GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 28, 2002, 16:44:30 ; Search time 1502.78 Seconds
(without alignments)
16229.254 Million cell updates/sec

Title: US-09-645-337-1
Perfect score: 1807
Sequence: 1 agagagcagctcccttcccc.....atgatgatatgatatgacaa 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	506	28.0	1627	11	BC013141	BC013141 Homo sapi
c	2	484.2	26.8	774	12	BH600970	BH600970 BOGCN27TR
	3	460.8	25.5	819	10	BE558877	BE558877 HV_CEb002
	4	452.2	25.0	917	10	BI118849	BI118849 EST237 Di
	5	427.4	23.7	760	10	BG596023	BG596023 EST494689
	6	423	23.4	749	10	BG594704	BG594704 EST493382
	7	422.4	23.4	535	10	W43397	W43397 22774 CD4-1
c	8	418.8	23.2	457	9	AI994124	AI994124 701499210
	9	412.6	22.8	999	9	AL518877	AL518877 AL518877
	10	408.6	22.6	775	10	BG368908	BG368908 HVSMEi002
	11	400.6	22.2	820	9	BE053804	BE053804 GA__Ea003
	12	398.8	22.1	919	9	AL531607	AL531607 AL531607
	13	393.8	21.8	656	9	AW396525	AW396525 sh28h04.y
	14	391.2	21.6	597	9	AW308961	AW308961 sf92d04.y
	15	386.4	21.4	847	9	AL558916	AL558916 AL558916
	16	385.6	21.3	1014	9	BB611086	BB611086 BB611086
	17	384.4	21.3	624	10	BF632187	BF632187 NF029B11D
c	18	380.6	21.1	1030	12	CNS06MLI	AL405628 T3 end of
	19	378.8	21.0	613	10	BG521741	BG521741 13-95 Ste
	20	378.8	21.0	986	9	AL536160	AL536160 AL536160
	21	377.6	20.9	696	10	BG597901	BG597901 EST496579
	22	377.2	20.9	437	10	R90034	R90034 16389 Lambd
	23	374.8	20.7	898	9	AL541430	AL541430 AL541430
	24	374.2	20.7	988	10	BM468325	BM468325 AGENCOURT
	25	373.4	20.7	882	10	BG367319	BG367319 HVSMEi001
	26	372.8	20.6	977	12	CNS072YS	AL426842 clone BA0
	27	372	20.6	905	10	BI855025	BI855025 603382030
	28	372	20.6	1007	10	BM452740	BM452740 AGENCOURT
	29	369.6	20.5	637	9	AW686294	AW686294 NF036C08N
	30	365.2	20.2	765	9	AJ394397	AJ394397 AJ394397
	31	363.4	20.1	910	10	BG756668	BG756668 602715527
	32	361	20.0	636	10	BI075248	BI075248 IP1_17_H0
	33	360	19.9	737	10	BI177518	BI177518 EST518463
	34	353.8	19.6	910	9	AL546378	AL546378 AL546378
	35	353.2	19.5	998	9	AL555361	AL555361 AL555361
	36	352.8	19.5	1034	10	BM461585	BM461585 AGENCOURT
	37	349.2	19.3	650	10	BG598526	BG598526 EST503426
	38	347	19.2	890	9	AL527966	AL527966 AL527966
	39	346	19.1	603	10	BE517647	BE517647 WHE0801_F
	40	345	19.1	345	9	AV536771	AV536771 AV536771
	41	344.8	19.1	528	9	AI736699	AI736699 sb32d06.y
	42	343.2	19.0	722	10	BG914796	BG914796 602813602
	43	343.2	19.0	871	10	BE747628	BE747628 601577573
	44	342.2	18.9	643	9	AW616696	AW616696 EST323107
	45	342.2	18.9	878	10	BF631408	BF631408 HVSMEb001

ALIGNMENTS

RESULT 1
BC013141
LOCUS BC013141 1627 bp mRNA linear HTC 29-AUG-2001
DEFINITION Homo sapiens, clone IMAGE:4475960, mRNA.
ACCESSION BC013141
VERSION BC013141.1 GI:15341911
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1627)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalon@bcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W., Muzny,D.M., Gibbs,R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 28 Row: d Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557640
This clone has the following problem: incomplete processing.
FEATURES Location/Qualifiers

```
     source          1. .1627
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4475960"
                     /tissue_type="Prostate, adenocarcinoma."
                     /clone_lib="NIH_MGC_91"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
BASE COUNT      479 a    321 c    401 g    426 t
ORIGIN


  Query Match             28.0%;  Score 506;  DB 11;  Length 1627;
  Best Local Similarity   63.3%;  Pred. No. 2e-117;
  Matches  776;  Conservative    0;  Mismatches  450;  Indels    0;  Gaps    0;

Qy    138 ggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagt 197
          |||   ||      | |||  |  |  |||||| | | |  |  || |  || |  |||||
Db    240 GGTGGCGGCGGTGGCCGGGGAGCCCATGGCGTACAGTCAAGGAGGCGGCAAAAAAAAAGT 299

Qy    198 ttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaa 257
          || || | ||| |||  |||  | || ||||| || ||||| || ||||||||||||||||
Db    300 CTGCTACTACTACGACGGTGATATTGGAAATTATTATTATGGACAGGGTCATCCCATGAA 359

Qy    258 gccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatat 317
          ||| ||| | ||||||||||||||||  | | ||    | || ||  |        | ||
Db    360 GCCTCATAGAATCCGCATGACCCATAACTTGCTGTTAAATTATGGCTTATACAGAAAAAT 419

Qy    318 gcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacga 377
          | |  |    | ||||      |||   |     ||  |       |  |||   || ||
Db    420 GGAAATATATAGGCCCCATAAAGCCACTGCCGAAGAAATGACAAAATATCACAGTGATGA 479

Qy    378 ctatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaact 437
          |||  ||   ||||| ||    || |  || || | |  |   ||  | | |   ||  |
Db    480 GTATATCAAATTTCTACGGTCAATAAGACCAGATAACATGTCTGAGTATAGTAAGCAGAT 539

Qy    438 taagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgcca 497
            || | || |||||||| ||||| ||||| || ||||| || || | |   |||||  ||
Db    540 GCAGAGATTTAATGTTGGAGAAGATTGTCCAGTGTTTGATGGACTCTTTGAGTTTTGTCA 599

Qy    498 gacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatat 557
          |  ||   |||| || || |||| |||  |||| ||| | ||||     |     |||||
Db    600 GCTCTCAACTGGCGGTTCAGTTGCTGGAGCTGTGAAGTTAAACCGACAACAGACTGATAT 659

Qy    558 tgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctg 617
          ||  | || |||||||| ||  | ||||| ||||||||  |  || || || ||  |||||
Db    660 GGCTGTTAATTGGGCTGGAGGATTACATCATGCTAAGAAATCAGAAGCATCAGGATTCTG 719

Qy    618 ttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttct 677
          ||||||  |||||||| ||  | ||  |||||  ||  | || ||| | ||| || | ||  |
Db    720 TTACGTTAATGATATTGTGCTTGCCATCCTTGAATTACTAAAGTATCATCAGAGAGTCTT 779

Qy    678 ttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactga 737
          |||  | |||||  |||||  || || ||  ||||| || || ||  || |||||| | || ||
Db    780 ATATATTGATATAGATATTCATCATGGTGATGGTGTTGAAGAAGCTTTTTATACAACAGA 839

Qy    738 cagggttatgactgtctcgtttcataaatttggtgattacttttcccggtacaggtcacat 797
            || ||||| || || || |||||||  ||| || ||||||||  || |||||   || |
Db    840 TCGTGTAATGACGGTATCATTCCATAAATATGGGGAATACTTTCCTGGCACAGGAGACTT 899

Qy    798 tcaggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatgg 857
            ||||||  |||  |||  |  ||  || |||||| || ||||| |  ||| ||   |||||
Db    900 GAGGGATATTGGTGCTGGAAAAGGCAAATACTATGCTGTCAATTTTCCAATGAGAGATGG 959

Qy    858 aatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaat 917
          || |||||||||   |||   | | |||| ||||| || ||    || || ||||| ||
Db    960 TATAGATGATGAGTCATATGGGCAGATATTTAAGCCTATTATCTCAAAGGTGATGGAGAT 1019

Qy    918 tttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggtt 977
          |  | |||  | ||||||||||| || |||||||| |||||  |||||||  ||| |  |
Db   1020 GTATCAACCTAGTGCTGTGGTATTACAGTGTGGTGCAGACTCATTATCTGGTGATAGACT 1079

Qy    978 aggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgtt 1037
          |||||  |||||||||  || ||||||||||||||  | || ||  || || | | | ||
Db   1080 GGGTTGTTTCAATCTAACAGTCAAAGGTCATGCTAAATGTGTAGAAGTTGTAAAAACTTT 1139

Qy    1038 caatgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttg 1097
           ||  | ||  ||||| |  | || || ||||| ||||| ||||| |||||||| || ||
Db    1140 TAACTTACCATTACTGATGCTTGGAGGAGGTGGCTACACAATCCGTAATGTTGCTCGATG 1199

Qy    1098 ctggtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagca 1157
           |||   || ||||||| |||||| ||||    |||  ||    |  || ||||  |  |
Db    1200 TTGGACATATGAGACTGCAGTTGCCCTTGATTGTGAGATTCCCAATGAGTTGCCATATAA 1259

Qy    1158 tgaatattatgaatactttggtccagactatacacttcacgttgctccaagtaacatgga 1217
           ||| || | ||| || ||||| |||||||  | ||| ||  ||  |||    ||||||
Db    1260 TGATTACTTTGAGTATTTTGGACCAGACTTCAAACTGCATATTAGTCCTTCAAACATGAC 1319

Qy    1218 aaataagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctc 1277
           |||  ||||  |||    |     | ||| ||||     |     |  |  | ||| |
Db    1320 AAACCAGAACACTCCAGAATATATGGAAAAGATAAAACAGCGTTTGTTTGAAAATTTGCG 1379

Qy    1278 taagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcc 1337
           | | | |  ||||| ||  |||| |   | || |  |  |||   ||| | |    |
Db    1380 CATGTTACCTCATGCACCTGGTGTCCAGATGCAAGCTATTCCAGAAGATGCTGTTCATGA 1439

Qy    1338 cgaggttgatgaagaccaagaagatg 1363
           ||   ||  || ||  | | ||| |
Db    1440 AGACAGTGTAGATGAAGATGGAGAAG 1465


RESULT   2
BH600970/c
LOCUS       BH600970                 774 bp    DNA     linear   GSS 15-DEC-2001
DEFINITION  BOGCN27TR BOGC Brassica oleracea genomic clone BOGCN27, DNA
            sequence.
ACCESSION   BH600970
VERSION     BH600970.1  GI:17853416
KEYWORDS    GSS.
SOURCE      Brassica oleracea.
  ORGANISM  Brassica oleracea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 774)
  AUTHORS   Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
  TITLE     Whole genome shotgun sequencing of Brassica oleracea
  JOURNAL   Unpublished (2001)
COMMENT     Other_GSSs: BOGCN27TF
            Contact: Chris Town
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA.
            Tel: 301-838-3523
            Fax: 301-838-0208
            Email: cdtown@tigr.org
            DNA is from a doubled haploid provided by Tom Osborn.
            Seq primer: TR
            Class: sheared ends.
FEATURES             Location/Qualifiers
     source          1. .774
                     /organism="Brassica oleracea"
                     /strain="TO1000DH3"
                     /db_xref="taxon:3712"
                     /clone="BOGCN27"
                     /clone_lib="BOGC"
                     /note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT      222 a    194 c    174 g    184 t
ORIGIN


  Query Match             26.8%;  Score 484.2;  DB 12;  Length 774;
  Best Local Similarity   85.4%;  Pred. No. 5.2e-112;
  Matches  551;  Conservative    0;  Mismatches   93;  Indels    1;  Gaps    1;

Qy    699 ccacggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgtt 758
```

```
Db    774 CCACGGGGACGGAGTGGAGGAGGCGTTTTACGCTACCGACAGGGTCATGACCGTCTCATT 715

Qy    759 tcataaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtag 818
Db    714 CCACAAATTCGGAGATTACTTTCCCGGTACAGGCCACATACAAGACATTGGCTACGGAAG 655

Qy    819 cggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatca 878
Db    654 TGGAAAGTACTATTCTCTCAACGTACCACTGGACGATGGAA-CGACGACGAGAGCTATCA 596

Qy    879 tctgttattcaagcccatcatggggaaagttatggaaatttccgaccaggggctgtggt 938
Db    595 TCTCTTGTTCAAACCCATCATGGGGAAAGTGATGGAAATTTTCAAACCAGGTGCGGTCGT 536

Qy    939 attgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaat 998
Db    535 ACTGCAGTGCGGCGCTGATTCCTTGTCTGGCGATAGGCTCGGATGCTTTAACCTCTCCAT 476

Qy    999 caaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctctt 1058
Db    475 CAAAGGCCACGCCGAGTGTGTCAAGTTCATGAGATCTTTCAACGTTCCGTTACTGCTCCT 416

Qy   1059 gggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagt 1118
Db    415 TGGTGGTGGTGGTTACACTATCCGCAACGTTGCGCGTTGCTGGTGCTACGAGACTGGAGT 356

Qy   1119 tgcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttgg 1178
Db    355 TGCGCTTGGGATCGAAGTCGAGGACAAGATGCCGGAGCATGAGTACTATGAGTACTTTGG 296

Qy   1179 tccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagat 1238
Db    295 TCCGGACTATACGCTTCACGTTGCTCCGAGTAACATGGAGAACAAGAACTCTCGGCAGAT 236

Qy   1239 gcttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaag 1298
Db    235 GCTTGAAGTGATTCGTAATGACTTGCTCCAGAATCTCTCGAAGCTTCAGCATGCTCCTAG 176

Qy   1299 tgtaccatttcaggaaagaccacctgatacagagactcccgaggt 1343
Db    175 TGTGCCGTTTCAGGAAAGGCCGCCAGATACTGAGACTCCTGAGGT 131
```

```
RESULT   3
BE558877
LOCUS       BE558877                819 bp    mRNA    linear   EST 23-OCT-2001
DEFINITION  HV_CEb0020K09f Hordeum vulgare seedling green leaf EST library
            HVcDNA0005 (Blumeria challenged) Hordeum vulgare cDNA clone
            HV_CEb0020K09f, mRNA sequence.
ACCESSION   BE558877
VERSION     BE558877.1  GI:9823367
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 819)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Wei,F., Begum,D.,
            Frisch,D., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi
            ,D.W., Fenton,R.D., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Blumeria infected incompatible (Mla6) seedling
            leaf cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 513
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 662.
FEATURES             Location/Qualifiers
     source          1. .819
                     /organism="Hordeum vulgare"
                     /cultivar="CI16151 (Mla6)"
                     /db_xref="taxon:4513"
                     /clone="HV_CEb0020K09f"
                     /clone_lib="Hordeum vulgare seedling green leaf EST
                     library HVcDNA0005 (Blumeria challenged)"
                     /tissue_type="seedling green leaf"
                     /lab_host="SOLR"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xhol;
                     C.I. 16151 (Mla6) plants were greenhouse grown in the R
                     Wise lab at Iowa State University, Ames, IA; 7 day old
                     green seedlings were challenged with isolate 5874 (AvrMla6
                     ) of Blumeria graminis f. sp. hordei, and leaves were
                     harvested 20 and 24 hr post-inoculation and snap frozen;
                     uninoculated leaves were harvested 20 hr post-inoculation
                     (Wei, Wise). In the TJ Close lab at the University of
                     California, Riverside, total RNA was prepared from each
                     sample pool, equal quantities of all three RNA pools were
                     combined, poly(A) RNA was purified from the mixture, one
                     primary unamplified cDNA library was made, and 1 million
                     pfu were in vivo excised to give pBluescript SK(-) cDNA
                     phagemids (Choi, Close). Phagemids were plated and picked
                     at the Clemson University Genomics Institute (CUGI) (Begum
                     , Palmer, Frisch, Atkins and Wing). Plasmid DNA
                     preparations, DNA sequencing and sequence analysis were
                     performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
                     , Rambo, Main). The sequence has been trimmed to remove
                     vector sequence and contains a minimum of 100 bases of
                     phred value 20 or above. For more details on library
                     preparation and sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT      173 a    239 c    216 g    191 t
ORIGIN
```

Query Match 25.5%; Score 460.8; DB 10; Length 819;
Best Local Similarity 74.9%; Pred. No. 4.7e-106;
Matches 617; Conservative 0; Mismatches 202; Indels 5; Gaps 3;

```
Qy    254 tgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagc 313
Db      1 TGAAGCCGCACCGCATCCGCATGACCCACTCGCTGCTGGCGCAGTACGGCCTCCTCGACC 60

Qy    314 atatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccg 373
Db     61 AGATGCAGGTGCTGCGCCCCAACCCCGCCCGCGACCGCGACCTCTGCCGCTTCCACGCCG 120

Qy    374 acgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgcc 433
Db    121 ACGACTACATCTCCTTCCTCCGCTCCGTCACGCCCGAGACGCAGCAGGACCAGATCCGGG 180

Qy    434 aacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttt 493
Db    181 CCCTCAAGCGCTTCAACGTCGGCGAGGACTGCCCCGTCTTCGACGGCCTCTACAGCTTCT 240

Qy    494 gccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcg 553
Db    241 GCCAGACCTACGCGGGGGCCTCCGTCGGCGGTGCCGTCAAGCTCAACCACGGTCT---CG 297

Qy    554 atattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggct 613
Db    298 ACATCGCCATCAACTGGTCAGGGGGCCTGCACCATGCTAAGAAGTGCGAGGCCTCAGGCT 357
```

```
Qy    614 tctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtg 673
          |||| |||||||| || |||||  | || ||||| |||||||| ||||| ||  |||| |
Db    358 TCTGCTACGTCAACGACATCGTACTCGCCATCCTCGAGCTCCTCAAGCATCACCAGCGAG 417

Qy    674 ttctttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgcta 733
          |||| ||||| ||||||||||||||||| || || ||||| |||||||||| ||| | |
Db    418 TTCTATATGTTGATATTGATATCCACCATGGTGACGGAGTTGAGGAGGCATTCTATACAA 477

Qy    734 ctgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtc 793
          | |||||||||||| ||||| || || || ||||| ||||| || || || |||||
Db    478 CAGACAGGGTTATGACGGTCTCATTCCACAAGTTTGGGGATTATTTCCCAGGCACAGGGG 537

Qy    794 acattcaggatataggttatggtagcggaaagtactattctctcaatgtaccactggatg 853
          |  | |  ||||| || ||  |  |||||||| || || || ||||| ||  ||||||
Db    538 ATGTCCGTGATATTGGGCATTCAAAAGGAAAGTATTACTCCCTTAATGTCCCTTTGGATG 597

Qy    854 atggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatgg 913
          |||| || ||||| || ||||| ||      |  | ||||| || || |   |||||||||
Db    598 ATGGGATTGATGAAGAAAGCTACCAGTCCCTTGTTAAGCCTATAAT-GCCCAAGTTATGG 656

Qy    914 aaatttttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatc 973
          || |||||| ||| || ||  | |   |  |||| | || |  | | ||||||| |||
Db    657 AAGTTTTCCAACCTGGTGCAATTGGTCTTTAATGCAGCGCCGGTTTCGTATCTGGAGATA 716

Qy    974 ggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagat 1033
          |||| || |||||||||||||| |||| | ||||||| || ||||| || |  ||||
Db    717 GGTTGGGGTGCTTCAATCTTTCGATCAGAAGTCATGCCGAATGCGTGAAGTACATGAAGG 776

Qy   1034 cgttcaatgttcccctactgctcttgggtggtggtggttacact 1077
          | ||| ||| |||  |  || | || || || || || || |||
Db    777 CTTTC-ATGGTCCATTGGTGGTTTTTGGCGGAGGGGGGTATACT 819


RESULT  4
BI118849
LOCUS       BI118849                 917 bp    mRNA    linear   EST 26-JUN-2001
DEFINITION  EST237 Differentially expressed cDNA libraries of BTH-treated/blast
            fungus-infected rice leaf tissues Oryza sativa cDNA clone HNBN-w14,
            mRNA sequence.
ACCESSION   BI118849
VERSION     BI118849.1  GI:14571481
KEYWORDS    EST.
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   Song,F.M. and Goodman,R.M.
  TITLE     Molecular cloning of differentially expressed cDNAs associated with
            systemic acquired resistance in rice
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Song, F.M.; Goodman, R.M.
            Department of Plant Pathology
            University of Wisconsin-Madison
            Rm 689, Russell Laboratories, 1630 Linden Drive, Madison, WI 53706,
            USA
            Tel: 608 262 9162
            Fax: 608 262 8643
            Email: fmsong@plantpath.wisc.edu; rgoodman@facstaff.wisc.edu
            BTH-downregulated
            Insert Length: 917   Std Error: 0.00
            Seq primer: M13 Forward and reverse.
FEATURES             Location/Qualifiers
     source          1. .917
                     /organism="Oryza sativa"
                     /cultivar="Yuanfengzao (a susceptible variety to rice
                     blast disease)"
                     /db_xref="taxon:4530"
                     /clone="HNBN-w14"
                     /clone_lib="Differentially expressed cDNA libraries of
                     BTH-treated/blast fungus-infected rice leaf tissues"
                     /tissue_type="Seedling leaves"
                     /dev_stage="Three-week-old rice seedlings pretreated with
                     BTH and/or inoculated with Magnaporthe grisea"
                     /note="Vector: pGEM-T Easy; Rice seedlings were pretreated
                     by foliar spraying with 0.3 mM  benzothiadiazole (BTH) and
                     then inoculated with Magnaporthe grisea three  days after
                     BTH treatment. Leaf samples were collected twenty-four
                     hour  after inoculation and mRNAs were used for
                     construction of the  differentially expressed cDNA
                     libraires by suppression subtractive  hybridization."
BASE COUNT      182 a    278 c    255 g    200 t      2 others
ORIGIN


  Query Match             25.0%;  Score 452.2;  DB 10;  Length 917;
  Best Local Similarity   74.4%;  Pred. No. 7.5e-104;
  Matches  688;  Conservative    1;  Mismatches  224;  Indels   12;  Gaps    9;

Qy    250 cccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcctt 309
          || |||||||||| |||||||| |||||||| || || ||||| |||||||| |||||
Db      2 CCGATGAAGCCCCACCGCATCCGGATGACCCACGCGCTGCTCGCCCACTACGGCCTCCTC 61

Qy    310 cagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccac 369
          | || |||||||| |||||||||  ||| || ||||| || || ||||||||||||||||
Db     62 GACCAGATGCAGGTGCTCAAGCCCCACCCGGCGCGCGACCGCGACCTCTGCCGCTTCCAC 121

Qy    370 gccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagatt 429
          |||||||||||| ||| | || ||||||  | | || || || |||||||| || |||||
Db    122 GCCGACGACTACGTCGCCTTCCTCCGCTCCGTCACGCCGGAGACCCAGCAGGACCAGATC 181

Qy    430 cgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattcc 489
          ||    || |||||||||| || || || ||||| |||||||| |||||||| ||   |
Db    182 CGGGCGCTCAAGCGCTTCAACGTCGGCGAGGACTGCCCCGTCTTCGACGGCCTCTACAGC 241

Qy    490 ttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctc 549
          || |||||||||| || || ||||| || || ||| | |||||||| ||||||||||
Db    242 TTCTGCCAGACCTACGCCGGGGGATCCGTCGGCGGCGCCGTCAAGCTCAACCACGGCC-- 299

Qy    550 tgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctct 609
            ||| || |||||||||||||| || || ||||| ||||| ||||||||||||||||||
Db    300 -ACGACATCGCCATCAACTGGGCCGGCGG-CTCCACCACGCCAAGAAGTGCGAGGCCTCC 357

Qy    610 ggcttctgttacgtcaatgatatcgtcttagctatcctagaagctccttaagcagcatgag 669
          |||||||| |||||||| || |||||| | || ||||| |||||||| ||  | ||  ||
Db    358 GGCTTCTGCTACGTCAACGACATCGTC-TCGCCATCCTCGAGCTCCTCAAATACCACCAG 416

Qy    670 cgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcattttat 729
          |||||||| ||||| || || |||||||||| |||||||| ||||||||||  ||||
Db    417 CGTGTTCTCTATGTGGACATCGATATCCACCATGGGGATGGTGTGGAGGAGGC-GTTTAC 475

Qy    730 gctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtaca 789
            | || |||||||| ||||| |||||||| || || ||||| :|||| || || || ||
Db    476 ACGACGGACAGGGTRATGACGGTCTCGTTCCACAAGTTTGGGRATTATTTCCCGGGGACC 535

Qy    790 ggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccactg 849
          ||  ||||||  ||||| ||  |     |  || ||||| || ||||| ||    ||  |
Db    536 GGGGACATTCGCGATATTGGGCA-CTCAAAGGGAAGTATTACTCTCTGAAATGTCCCGTT 594

Qy    850 gatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatgggqaaagtt 909
          || ||  | || ||  || ||||||| ||   |||  |||||||  |||||| ||  |
Db    595 GACGA-CGTATTGACAATAAGAGCTACCAGTCGTT-GTCAAGCCGTTCATGGGAAAGTTA 652

Qy    910 atggaaattttccgaccaggggctgtggtattgcaatgtggt--gctgactccctatctg 967
          ||| |  ||||    || ||  | || ||  | |||  ||||  |    || |      |
Db    653 ATGAAGGTTTTTGCCCCTGGCCCGGTTGTTGTTCAAGTTGGTCGGACTTCTTCTGTCAAG 712

Qy    968 gggatcggttaggttgcttcaa-tctttcaatcaaaggtcatgctgagtgcgtcaaattt 1026
          | |||||||| || |||||||  || || |||  || || || || || || | |||
Db    713 GTGATCGGTTGGGCTGCTTCAACCCTATCGATCAGGGGCCACGCGGAATGTGTGAGATTC 772

Qy   1027 atgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaat 1086
          ||||| || ||||| || || ||  ||||  | |||||||||||| || || ||  | |||
Db    773 ATGAGGTCCTTCAACGTCCCGCTGTTGCTGCTTGGTGGTGGTGGGTATACCATAAGAAAT 832
```

```
Qy   1087 gttgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaagacaag 1146
          ||||| || || |||||||| ||||| |||||||||||||||   ||| |    ||||||
Db    833 GTTGCACGGTGTTGGTGCTATGAGACAGGAGTTGCACTTGGTCATGAGCTCACCGACAAG 892

Qy   1147 atgccggagcatgaatattatgaat 1171
          |||||     |||| |||| ||| |
Db    893 ATGCCTCCAAATGAGTATTTTGAGT 917


RESULT   5
BG596023
LOCUS       BG596023                760 bp    mRNA    linear   EST 12-APR-2001
DEFINITION  EST494689 cSTS Solanum tuberosum cDNA clone cSTS12O24 5' sequence,
            mRNA sequence.
ACCESSION   BG596023
VERSION     BG596023.1  GI:13614151
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
  TITLE     Generations of ESTs from sprouting potato eyes
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cathy Ronning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cdna@resgen.com
            Seq primer: M13F-R.
FEATURES             Location/Qualifiers
     source          1. .760
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="cSTS12O24"
                     /clone_lib="cSTS"
                     /tissue_type="sprouting eyes from tubers"
                     /dev_stage="12-14 weeks post harvest"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
                     taken from tubers. The tubers were incubated at 26C in the
                     dark for 2-3 weeks prior to sprouting. The eyes were
                     frozen in liquid nitrogen immediately upon removal from
                     tubers."
BASE COUNT      198 a    138 c    194 g    230 t
ORIGIN


  Query Match             23.7%;  Score 427.4;  DB 10;  Length 760;
  Best Local Similarity   76.0%;  Pred. No. 1.4e-97;
  Matches  527;  Conservative    0;  Mismatches  166;  Indels    0;  Gaps    0;

Qy    138 ggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagt 197
          ||  ||||||  ||| || || ||  | || || ||  | |||||   |||||| |||||
Db     68 GGACATGGATGTTGGAGGAAACTCCTTAGCATCTGGTGCAGATGGAAAGAAGAGAAAAGT 127

Qy    198 ttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaa 257
          ||||||||||||| ||||| || |||||||||| ||||| |||| || || |||||
Db    128 AAGTTATTTCTATGATCCTGAAGTTGGCAATTACTATTATGGACAAGGCCACCCAATGAA 187

Qy    258 gccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatat 317
          ||| ||||| || || ||||| ||||| || || || ||||| |||||  | || |||||
Db    188 GCCACATCGAATGCGGATGACACATGCTCTTCTTGCCCACTATGGTCTATTACAACATAT 247

Qy    318 gcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacga 377
          ||| ||||| |||||   ||||||  | ||    |||||||| | || || || || ||
Db    248 GCATGTTCTGAAGCCAAACCCTGCTAGAGATAAGGATCTCTGCAGGTTTCATGCTGATGA 307

Qy    378 ctatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaact 437
          |||| |  ||||  |  | ||| | ||||| ||||| |||||||||||| |  | || ||
Db    308 TTATATTGCTTTCTTGAGAAGCGTAACCCCAGAAACACAGCAAGATCAGCTGAGGCAGCT 367

Qy    438 taagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgcca 497
          ||| | ||||| ||||| || |||||||| || ||||| || |||||||||| |||||
Db    368 GAAGAGATTCAACGTTGGGGAGGACTGTCCAGTTTTTGATGGTCTTTATTCCTTCTGCCA 427

Qy    498 gacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatat 557
          |||||||||||||| ||||||||||  | || ||| | ||||| || | ||| |||||
Db    428 AACCTATGCTGGAGGTTCTGTTGGTGGGGCAGTGAAGTTGAACCATGGACACTGTGATAT 487

Qy    558 tgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctg 617
          ||  | || ||||||||||||  | ||||| || ||||| || || || ||||| |||||
Db    488 CGCTGTAAATTGGGCTGGTGGATTGCATCATGCCAAGAAATGTGAAGCATCTGGTTTCTG 547

Qy    618 ttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttct 677
          || || || |||||||| || ||||| || ||||| || ||    |||||| | ||  |
Db    548 CTATGTGAACGATATCGTGTTGGCTATTCTGGAGCTACTCAAAGTTCATGAGAGAGTGTT 607

Qy    678 ttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactga 737
          ||||| |||||||||||||| || ||||| ||||| || || || ||| | || ||
Db    608 GTATGTGGATATTGATATCCATCATGGTGATGGCGTGGAAGAAGCCTTCTATACCACAGA 667

Qy    738 cagggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacat 797
          ||||| ||||| || || |||||||||||||| |||||||||| |||||||  |  |
Db    668 TAGGGTCATGACAGTTTCCTTTCATAAATTTGGAGATTACTTTCCTGGTACAGGGGATGT 727

Qy    798 tcaggatataggttatggtagcggaaagtacta 830
          |  ||||| |||||||| || ||||| |||||
Db    728 ACGTGATATTGGTTATGGAAGTGGAAAATACTA 760


RESULT   6
BG594704
LOCUS       BG594704                749 bp    mRNA    linear   EST 12-APR-2001
DEFINITION  EST493382 cSTS Solanum tuberosum cDNA clone cSTS8A2 5' sequence,
            mRNA sequence.
ACCESSION   BG594704
VERSION     BG594704.1  GI:13612844
KEYWORDS    EST.
SOURCE      potato.
  ORGANISM  Solanum tuberosum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
            Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
  TITLE     Generations of ESTs from sprouting potato eyes
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cathy Ronning
            The Institute for Genomic Research
            For clone info: please contact Research Genetics, Libraries
            Division tel 1-800-711-6195, email cdna@resgen.com
            Seq primer: M13F-R.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Solanum tuberosum"
                     /cultivar="Kennebec"
                     /db_xref="taxon:4113"
                     /clone="cSTS8A2"
                     /clone_lib="cSTS"
                     /tissue_type="sprouting eyes from tubers"
                     /dev_stage="12-14 weeks post harvest"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
                     taken from tubers. The tubers were incubated at 26C in the
                     dark for 2-3 weeks prior to sprouting. The eyes were
                     frozen in liquid nitrogen immediately upon removal from
                     tubers."
BASE COUNT      196 a    137 c    190 g    226 t
ORIGIN
```

```
  Query Match             23.4%;  Score 423;  DB 10;  Length 749;
  Best Local Similarity   76.0%;  Pred. No. 1.8e-96;
  Matches  522;  Conservative    0;  Mismatches  165;  Indels    0;  Gaps    0;

Qy    138 ggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagt 197
          ||  ||||||  ||| || || ||  | || || ||  | |||||  ||||||| |||||
Db     55 GGACATGGATGTTGGAGGAAACTCCTTAGCATCTGGTGCAGATGGAAAGAAGAGAAAAGT 114

Qy    198 ttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaa 257
            |||||||||||||| ||||| || |||||||||||| ||||| ||||| || || |||||
Db    115 AAGTTATTTCTATGATCCTGAAGTTGGCAATTACTATTATGGACAAGGCCACCCAATGAA 174

Qy    258 gccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatat 317
          ||| ||||| || || ||||| ||||| || || || ||||| |||||  | || |||||
Db    175 GCCACATCGAATGCGGATGACACATGCTCTTCTTGCCCACTATGGTCTATTACAACATAT 234

Qy    318 gcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacga 377
          ||| ||||| |||||   ||||||  | ||    |||||||||| | || || || || ||
Db    235 GCATGTTCTGAAGCCAAACCCTGCTAGAGATAAGGATCTCTGCAGATTTCATGCTGATGA 294

Qy    378 ctatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaact 437
          ||| |  ||||  |  | ||| | ||||| ||||| ||||||||||||| |  | || ||
Db    295 TTATATTGCTTTCTTGAGAAGCGTAACCCCAGAAACACAGCAAGATCAGCTGAGGCAGCT 354

Qy    438 taagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgcca 497
          ||| | ||||| ||||| || |||||||| || ||||| || |||||||||| |||||
Db    355 GAAGAGATTCAACGTTGGGGAGGACTGTCCAGTTTTTGATGGTCTTTATTCCTTCTGCCA 414

Qy    498 gacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatat 557
          ||||||||||||||| ||||||||||  | || ||| | ||||| || | ||| |||||
Db    415 AACCTATGCTGGAGGTTCTGTTGGTGGGGCAGTGAAGTTGAACCATGGACACTGTGATAT 474

Qy    558 tgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctg 617
          ||  | || ||||||||||  | ||||| || ||||| || || || ||||| |||||
Db    475 CGCTGTAAATTGGGCTGGTGGATTGCATCATGCCAAGAAATGTGAAGCATCTGGTTTCTG 534

Qy    618 ttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttct 677
          || || || |||||||| || ||||| || ||||| || ||    |||||| | ||  |
Db    535 CTATGTGAACGATATCGTGTTGGCTATTCTGGAGCTACTCAAAGTTCATGAGAGAGTGTT 594

Qy    678 ttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactga 737
          ||||| |||||||||||||| || || ||||| ||||| || || || ||| | || ||
Db    595 GTATGTGGATATTGATATCCATCATGGTGATGGCGTGGAAGAAGCCTTCTATACCACAGA 654

Qy    738 cagggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacat 797
          ||||| ||||| || || |||||||||||||| |||||||||| ||||||||  |  |
Db    655 TAGGGTCATGACAGTTTCCTTTCATAAATTTGGAGATTACTTTCCTGGTACAGGGGATGT 714

Qy    798 tcaggatataggttatggtagcggaaa 824
          |  ||||| |||||||| || |||||
Db    715 ACGTGATATTGGTTATGGAAGTGGAAA 741


RESULT   7
W43397
LOCUS       W43397                   535 bp    mRNA    linear   EST 05-JAN-1998
DEFINITION  22774 CD4-15 Arabidopsis thaliana cDNA clone G11C3T7, mRNA
            sequence.
ACCESSION   W43397
VERSION     W43397.1  GI:2748746
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
            ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
            ,E. and Somerville,C.
  TITLE     Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
  JOURNAL   Plant Physiol. 106, 1241-1255 (1994)
  MEDLINE   95148729
COMMENT     On Jan 5, 1998 this sequence version replaced gi:1327865.
            Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University,Plant Biology Bldg.,E.
            Lansing,Mi
            Tel: 517-353-0854
            Fax: 517-353-9168
            Email: 22313tcn@ibm.cl.msu.edu
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .535
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="G11C3T7"
                     /clone_lib="CD4-15"
                     /tissue_type="seedling hypocotyl"
                     /dev_stage="3 day-old"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     EcoRI; Using 5 ug of polyadenylated mRNA from 3 day-old
                     Arabidopsis thaliana (Columbia) seedling hypocotyls as
                     template and oligo d(t) as primer, first strand synthesis
                     was catalyzed by Moloney murine leukemia virus reverse
                     transcriptase (Pharmacia).  Second-strand cDNA was made
                     using the procedure of Gubler and Hoffman (1983) except
                     that DNA ligase was omitted.  After the second strand
                     reaction, the ends of the cDNA were made blunt with Klenow
                     fragment and EcoRI/NotI adapters (Pharmacia) were ligated
                     to each end.  The cDNA was purified from unligated
                     adapters by spun-column chromatography using sephacryl
                     s-300 and size-fractionated on a 1% low melting point
                     mini-gel.  Size selected cDNAs (2 - 3 kb) were removed
                     from the gel using agarase (New England Biolabs),
                     phenol:choloroform extracted and precipitated using 0.3 M
                     NaOAc (pH 7)/ethanol.  A portion of each cDNA
                     size-fraction (0.1 ug) was co-precipitated with 1 ug of
                     lambdaZapII (StratageneEcoRI digested, dephosphorylated
                     arms and then ligated in a volume of 4 ul overnight. Each
                     ligation mix was packaged in vitro using Giqapack II gold
                     packaging extract (Stratagene).  We have determined that
                     although first strand cDNA synthesis was initiated using
                     dT, almost all of the cDNAs begin 8-10 bp from the poly-A
                     tail.  The reason for the loss of the poly-A tail is most
                     likely due to lower than anticipated nucleotide levels
                     during the Klenow repair of ragged ends before the
                     addition of linkers (3'-5' exo instead of 5'-3' pol).
                     When this library is used please reference the ABRC and:
                     Kieber, J.  et al. (1993) Cell 72:427-441."
BASE COUNT      127 a    143 c    133 g    122 t     10 others
ORIGIN


  Query Match             23.4%;  Score 422.4;  DB 10;  Length 535;
  Best Local Similarity   97.6%;  Pred. No. 2.2e-96;
  Matches  448;  Conservative    0;  Mismatches    9;  Indels    2;  Gaps    2;

Qy      2 gagagcagctcccttcccctcggcgaggaggaaggaagaagaaagccagagagagagaga 61
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAGAGCAGCTCCCTTCCCCTCGGCGAGGAGGAAGGAAGAAGAAAGCCAGAGAGAGAGAGA 60

Qy     62 gagatcatcgcagcttctcctccgaccatttgactgcgactgtgattacaacacaccgtt 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAGATCATCGCAGCTTCTCCTCCGACCATTTGACTGCGACTGTGATTACAACACACCGTT 120

Qy    122 gatcctacgaaaaagaggtaatggatactggcggcaattcgctggcgtccggacctgatg 181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GATCCTACGAAAAAGAGGTAATGGATACTGGCGGCAATTCGCTGGCGTCCGGACCTGATG 180

Qy    182 gtgtgaagaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggcc 241
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GTGTGAAGAGGAAAGTTTGTTATTTCTATGACCCTGAGGTCGGCAATTACTACTATGGCC 240

Qy    242 aaggtcatcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacg 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 AAGGTCATCCCATGAAGCCCCATCGCATCCGCATGACCCATGCCCTCCTCGCTCACTACG 300

Qy    302 gtctccttcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgcc 361
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    301 GTCTCCTTCAGCATATGCAGGTTCTCAAGCCCTTCCCTGCCCGCGAACGTGATCTCTGNC 360

Qy    362 gcttccacgccg-acgactatgtctcttttctccgcagcattacccctgaaacccagcaa 420
          | |||||||||| ||||||||||||||||||||||||||| |||||||| |||||||
Db    361 GNTTCCACGCCGNACGACTATGTCTCTTTTCTCCGCAGCATTAACCCTGAAAACCAGCAA 420

Qy    421 gatcagattcgccaacttaagcgcttcaatgttggtgaa 459
          |||||||||||| |||||||||||  |   |  |||||||||
Db    421 GATCAGATTCG-CAACTTAAGCGGTNAATGGTTGGTGAA 458


RESULT   8
AI994124/c
LOCUS       AI994124                 457 bp    mRNA    linear   EST 08-SEP-1999
DEFINITION  701499210 A. thaliana, Ohio State clone set Arabidopsis thaliana
            cDNA clone 701499210, mRNA sequence.
ACCESSION   AI994124
VERSION     AI994124.1  GI:5841029
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
            Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,
            Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R.
            , Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T.,
            Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L.,
            Turner,C., Krikorian,S., Elder,L. and Hanson,D.
  TITLE     Arabidopsis thaliana Gene Expression MicroArray
  JOURNAL   Unpublished (1999)
COMMENT     Contact: David Smoller, Ph.D.
            Genome Systems, Inc., a wholly owned subsidiary of Incyte
            Pharmaceuticals, Inc.
            4633 World Parkway Circle, St. Louis, MO 63134, USA
            Tel: 877-577-2733
            Fax: 314-427-3324
            Email: service@genomesystems.com.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="701499210"
                     /clone_lib="A. thaliana, Ohio State clone set"
                     /note="cDNA library was made from selected clones from the
                     Arabidopsis thaliana Ohio  State clone set."
BASE COUNT      101 a    132 c     61 g    140 t     23 others
ORIGIN


  Query Match             23.2%;  Score 418.8;  DB 9;  Length 457;
  Best Local Similarity   94.1%;  Pred. No. 1.7e-95;
  Matches  432;  Conservative    0;  Mismatches   25;  Indels    2;  Gaps    1;

Qy   1336 cccgaggttgatgaagaccaagaagatggggataaaagatgggatccggattcagacatg 1395
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    457 CCCGAGGTTGATGAAGACCAAGAAGATGGGGATAAAAGATGGGATCCGGATTCAGACATG 398

Qy   1396 gatgttgatgatgaccgtaaacctataccaagcagagtaaaaagagaagctgttgaacca 1455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    397 GATGTTGATGATGACCGTAAACCTATACCAAGCAGAGTAAAAAGAGAAGCTGTTGAACCA 338

Qy   1456 gatacaaaggacaaggatggactgaaaggaattatggagcgtggaaaaggttgtgaggtg 1515
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    337 GATACTAAGGACAAGGATGGACTGAAAGGAATTATGGAGCGTGGAAAAGGTTGTGAGGTG 278

Qy   1516 gaggtggatgagagtggaagcactaaggttacaggagtaaacccagtgggagtggaggaa 1575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    277 GAGGTGGATGAGAGTGGAAGCACTAAGGTTACAGGAGTAAACCCAGTGGGAGTGGAGGAA 218

Qy   1576 gcaagtgtgaaaaatggaagaggaaggaacaaacaagggtggggcggagcaggcgtttcct 1635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    217 GCAAGTGTGAAAATGGAAGAGGAAGGAACAAACAAGGGTGGGGCGGAGCAGGCGTTTCCT 158

Qy   1636 cctaaaacataagactcggagcttctaatttcttgctacttttttctgtctatcaaatgtt 1695
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    157 CCTAAAACATAAGACTCGGAGCTTCTAATTTCTTGCTACTTTTTCTGTCTATCAAATGTT 98

Qy   1696 gctagttaagtttctggagttgttgttgttgtaagcactcctctgttttagaggattgag 1755
          |||||||||                          ||||||||||||||||||||||||||||
Db     97 GCTAGTTAANNNNNNNNNNNNNNNNNNNNNNNNAAGCACTCCTCTGTTTTAGAGGATTGAG 38

Qy   1756 cacggatatgtatttattcgttgcatgtctgaatgatga 1794
          |||||  |||| ||||||||||||||||||||||||||
Db     37 CACGG--ATGTGTTTATTCGTTGCATGTCTGAATGATGA 1


RESULT   9
AL518877
LOCUS       AL518877                 999 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL518877 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA011YG11 5
            prime, mRNA sequence.
ACCESSION   AL518877
VERSION     AL518877.1  GI:12782370
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 999)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .999
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DA011YG11"
                     /clone_lib="LTI_NFL011_NBC1"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library is not normalized, but is the control for
                     the normalized libraries. Library was constructed by Life
                     Technologies.  Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      257 a    223 c    275 g    234 t     10 others
ORIGIN


  Query Match             22.8%;  Score 412.6;  DB 9;  Length 999;
  Best Local Similarity   66.0%;  Pred. No. 8.9e-94;
  Matches  620;  Conservative    4;  Mismatches  313;  Indels    2;  Gaps    2;
```

```
Qy     189 gaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtca 248
           ||||||||| ||||| | ||| |||   || || || |||||||| :|||| ||||| ||
Db      47 GAGGAAAGTCTGTTACTACTACGACGGGGATGTTGGAAATTACTATWATGGACAAGGCCA 106

Qy     249 tcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcct 308
           || |||||||| || || ||||||||||| ||:    | ||     |||| ||||||
Db     107 CCCAATGAAGCCTCACCGAATCCGCATGACTCAWAATTTGCTGCTCAACTATGGTCTCTA 166

Qy     309 tcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttcca 368
           |   | ||| |  |        ||   |  |||   |     ||  |    |   | |||
Db     167 CCGAAAAATGGAAATCTATCGCCCTCACAAAGCCAATGCTGAGGAGATGACCAAGTACCA 226

Qy     369 cgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagat 428
           |  ||| |||||  |    ||  | |||  |||    || || | |  |   ||  | |
Db     227 CAGCGATGACTACATTAAATTCTTGCGCTCCATCCGTCCAGATAACATGTCGGAGTACAG 286

Qy     429 tcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattc 488
           ||  |  || | ||||| |||||||| |||||||| || || || |||||  | |
Db     287 CAAGCAGATGCAGAGATTCAACGTTGGTGAGGACTGTCCAGTATTCGATGGCCTGTTTGA 346

Qy     489 cttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcct 548
           || || |||   | | |||| || ||||| |   |  |||| || |||||  |    |
Db     347 GTTCTGTCAGTTGTCTACTGGTGGTTCTGTGGCAAGTGCTGTGAAACTTAATAAGCAGCA 406

Qy     549 ctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctc 608
           ||  || ||  | || |||||||| || || || || || |||||||  |||||| ||
Db     407 GACGGACATCGCTGTGAATTGGGCTGGGGGCCTGCACCATGCAAAGAAGTCCGAGGCATC 466

Qy     609 tggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatga 668
           |||||||||||||||||||| |||||||| || ||||| || || || ||| | ||  |
Db     467 TGGCTTCTGTTACGTCAATGAAATCGTCTTGGCCATCCTGGAACTGCTAAAGTATCACCA 526

Qy     669 gcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttta 728
           | | || || ||  | || |||||||| ||||| || || || ||||| ||||| || ||
Db     527 GAGGGTGCTGTACATTGACATTGATATTCACCATGGTGACGGCGTGGAAGAGGCCTTCTA 586

Qy     729 tgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtac 788
           | || ||| |||| |||||||| || |||||||| | ||| || ||||| || || ||
Db     587 CACCACGGACCGGGTCATGACTGTGTCCTTTCATAAGTATGGAGAGTACTTCCCAGGAAC 646

Qy     789 aggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccact 848
           ||  || | | |||||  ||   ||| |  || |||||  ||| || | ||    || ||
Db     647 TGG-GACCTACGGGATAACGGGGCTGGCAAAGGCAAGTATTATGCTGTTAACTACCCGCT 705

Qy     849 ggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagt 908
           || || ||  |||| |||  |||| |     | ||||||  |||||   |||||
Db     706 CCGAGACGGGATTTATGACGAGTYCTATGAGGCCATTTTCAAGCCGGTCATGTYCAAAGT 765

Qy     909 tatggaaatttccgaccaggggctgtggtattgcaatgtggtgctgactccctatct-g 967
           |||||  || ||||  ||  | || |||||  || || |||||  | ||||||||||| |
Db     766 AATGGAGATGTTCCAGCCTAGTGCGGTGGTCTTACAGTGTGGCTCAGACTCCCTATCTGG 825

Qy     968 gggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttta 1027
           ||||||||||||||||||||||||  | |||||||| || ||  |||| ||  |||||
Db     826 GGGATCGGTTAGGTTGCTTCAATCTAACTATCAAAGGACACGCCAAGTGTGTKGAATTTG 885

Qy    1028 tgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaatg 1087
           | |      || ||  | ||  | |:| |  :||| || ||||||||||  | || || |
Db     886 TCAAGAGCTTTAACCTGCCTATKCYGATGCYGGGAGGCGGTGGTTACACAYTTCGTAACG 945

Qy    1088 ttgcccgttgctggtgctacgagactggagttgcacttg 1126
           ||||||| ||||||    | ||||  |  || || || |
Db     946 TTGCCCGGTGCTGGACAAATGAGAAAGCYGTGGCCCTGG 984


RESULT  10
BG368908
LOCUS       BG368908                 775 bp    mRNA    linear   EST 22-OCT-2001
DEFINITION  HVSME10021D24f Hordeum vulgare 20 DAP spike EST library HVcDNA0010
            (20 DAP) Hordeum vulgare cDNA clone HVSME10021D24f, mRNA sequence.
ACCESSION   BG368908
VERSION     BG368908.1  GI:13258009
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
            ,R.D., Close,S.J., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Morex 20 DAP spike cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 631
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence start: 9
            High quality sequence stop: 706.
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Hordeum vulgare"
                     /cultivar="Morex"
                     /db_xref="taxon:4513"
                     /clone="HVSMEi0021D24f"
                     /clone_lib="Hordeum vulgare 20 DAP spike EST library
                     HVcDNA0010 (20 DAP)"
                     /tissue_type="20 DAP spike"
                     /lab_host="SOLR"
                     /note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
                     Plants were grown in the greenhouse at the University of
                     California, Riverside (Fenton, SJ Close, TJ Close). Whole
                     spikes with awns trimmed were collected at 20 DAP (Fenton
                     ). Total RNA was prepared, poly(A) RNA was purified, one
                     primary unamplified cDNA library was made, and 1 million
                     pfu were in vivo excised to give pBluescript SK(-) cDNA
                     phagemids in the TJ Close lab at the University of
                     California, Riverside (Choi). Phagemids were plated and
                     picked at the Clemson University Genomics Institute (CUGI)
                     (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
                     preparations, DNA sequencing and sequence analysis were
                     performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
                     , Rambo, Main). The sequence has been trimmed to remove
                     vector sequence and contains a minimum of 100 bases of
                     phred value 20 or above. For more details on library
                     preparation and sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT      155 a    280 c    216 g    121 t      3 others
ORIGIN


  Query Match             22.6%;  Score 408.6;  DB 10;  Length 775;
  Best Local Similarity   78.4%;  Pred. No. 8.3e-93;
  Matches  502;  Conservative    0;  Mismatches  135;  Indels    3;  Gaps    1;

Qy     170 ccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgaggtcggcaatt 229
           ||||  | || ||   |||| |    || || || ||||| ||| | ||||| ||||| |
Db     125 CCGGCGCGGACGGCTCGAAGCGCCGCGTCTGCTACTTCTACGACGCGGAGGTGGGCAACT 184

Qy     230 actactatggccaaggtcatcccatgaagccccatcgcatccgcatgacccatgccctcc 289
           ||||||| || || || || || |||||||| || ||||||||||||||||| ||||| |
Db     185 ACTACTACGGGCAGGGCCACCCGATGAAGCCGCACCGCATCCGCATGACCCACGCCCTGC 244

Qy     290 tcgctcactacggtctccttcagcatatgcaggttctcaagccctccctgcccgcgaac 349
```

```
Db    245 TCGCCCACTACGGCCTCCTCGACGAGATGCAGGTGCTCAAGCCGCACCCCGCCCGCGACC 304

Qy    350 gtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagcattacccctg 409
Db    305 GCGACCTCTGCCGCTTCCACGCCGACGACTACGTCTCCTTCCTCCGCTCCGTCACCCCGG 364

Qy    410 aaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaagactgtcccg 469
Db    365 AGACGCAGCAGGACCAGATCCGCGCGCTCAAGCGCTTCAACGTCGGCGAGGACTGCCCCG 424

Qy    470 tctttgacggcctttattccttttgccagacctatgctggaggatctgttggtggctctg 529
Db    425 TCTTCGACGGCCTCTACAGCTTCTGCCAGACCTACGCCGGCGGCTCCGTCGGGGGCGCCG 484

Qy    530 tcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggtctccatcacg 589
Db    485 TCAAGCTCAACCACGGCCA---CGACATCGCAATCAACTGGGCCGGCGGCCTCCACCACG 541

Qy    590 ctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtcttagctatcctag 649
Db    542 CCAAGAAGTGCGAGGCCTCCGGCTTCTGCTACGTCAACGACATCGTCCTCGCCATCCTCG 601

Qy    650 agctccttaagcagcatgagcgtgttctttatgtcgatattgatatccaccacggggatg 709
Db    602 AGCTCCTCAAATACCACCAGCGTGTTCTGTATGTCGATATCGATATCCACCATGGGGACC 661

Qy    710 gagtggaggaggcattttatgctactgacagggttatgactgtctcgtttcataaatttg 769
Db    662 GCGTGGAGGAGGCGTTTTACACCACGGACAGGGTGATGACGGTCTCATTCCACAATTTTG 721

Qy    770 gtgattactttcccggtacaggtcacattcaggatatagg 809
Db    722 GGGATTATTNTCCCAGGACAGGAGACATTCGTGACGTTGG 761


RESULT  11
BE053804
LOCUS       BE053804                 820 bp    mRNA    linear   EST 07-MAR-2001
DEFINITION  GA__Ea0035C07f Gossypium arboreum 7-10 dpa fiber library Gossypium
            arboreum cDNA clone GA__Ea0035C07f, mRNA sequence.
ACCESSION   BE053804
VERSION     BE053804.2  GI:13246713
KEYWORDS    EST.
SOURCE      Gossypium arboreum.
  ORGANISM  Gossypium arboreum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE   1  (bases 1 to 820)
  AUTHORS   Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
            ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
  TITLE     An integrated analysis of the genetics, development, and evolution
            of the cotton fiber
  JOURNAL   Unpublished (2000)
COMMENT     On Jun 8, 2000 this sequence version replaced gi:8380860.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            High quality sequence stop: 818.
FEATURES             Location/Qualifiers
     source          1. .820
                     /organism="Gossypium arboreum"
                     /strain="AKA"
                     /cultivar="8400"
                     /db_xref="taxon:29729"
                     /clone="GA__Ea0035C07f"
                     /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
                     /tissue_type="Fibers isolated from bolls harvested 7-10
                     dpa"
                     /lab_host="E. coli"
                     /note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      257 a    169 c    190 g    204 t
ORIGIN


  Query Match             22.2%;  Score 400.6;  DB 9;  Length 820;
  Best Local Similarity   76.4%;  Pred. No. 9.2e-91;
  Matches  519;  Conservative    0;  Mismatches  154;  Indels    6;  Gaps    2;

Qy    937 gtattgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttca 996
Db      1 GTCCTTCAATGTGGTGCTGACTCCCTATCTGGAGATCGGTTGGGGTGTTTCAATCTTTCA 60

Qy    997 atcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctc 1056
Db     61 ATCAAGGGTCATGCTGAATGTGTTAAATTTATGAGATCTTTCAATGTTCCATTGCTACTA 120

Qy   1057 ttgggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactgga 1116
Db    121 TTAGGTGGTGGTGGTTATACCATTCGGAATGTTGCTCGTTGTTGGTGCTATGAGACAGGA 180

Qy   1117 gttgcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatacttt 1176
Db    181 GTTGCACTCGGAACGGAAGTTGACGACCAGCTGCCTCAACATGAATATTATGAATATTTC 240

Qy   1177 ggtccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcag 1236
Db    241 GGTCCGGATTATACACTTCATGTTGCCCCTAGTAACATGGAAAACAAGAACACACGCCAA 300

Qy   1237 atgcttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctcca 1296
Db    301 TTACTTGAAGAAATACGTAATAATCTACTTGACAATCTCTCAAAGCTTCAGCATGCACCA 360

Qy   1297 agtgtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaagaccaa 1356
Db    361 AGTGTCCAATTTCAAGAAAGACCACCTGACACTGAGCTTCCCGAGGCCGATGAAGATCAA 420

Qy   1357 gaagatggggataaaagatgggatccggattcagacatggatgttgatgatgaccgtaaa 1416
Db    421 GATGATGCAGATGAAAGATGTGACCCGGATTCAGACATGGAGGTTGATGAGAAGCG---G 477

Qy   1417 cctataccaagcagagtaaaaagagaagctgttgaaccagatacaaaggacaaggatgga 1476
Db    478 CCAATCCCTAGCAGAGTAAAGAGAGAAGCAATTGAACCCGAGAGAAAAGATTCGGAGGCT 537

Qy   1477 ctgaaaggaattatggagcgtggaaaagg---ttgtgaggtggaggtggatgagagtgga 1533
Db    538 CCGAAAGGCACTGCAGAGCAAGCTAGAGGGATTCGACACATGGCAACAAATGAGACCCTA 597

Qy   1534 agcactaaggttacaggagtaaacccagtgggagtggaggaagcaagtgtgaaaatggaa 1593
Db    598 AGTACAAAGTCTGTAGATGTCAGGCCTATGGCTATCGATGAACCAACTGTGAAAGTCGAG 657

Qy   1594 gaggaaggaacaaacaagg 1612
Db    658 CAAGAAACATCGAACAAAG 676


RESULT  12
AL531607
LOCUS       AL531607                 919 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL531607 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM002YH15 5
            prime, mRNA sequence.
ACCESSION   AL531607
VERSION     AL531607.1  GI:12795100
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 919)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .919
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DM002YH15"
                     /clone_lib="LTI_NFL001_NBC4"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      233 a    208 c    253 g    223 t      2 others
ORIGIN
```

Query Match 22.1%; Score 398.8; DB 9; Length 919;
Best Local Similarity 66.5%; Pred. No. 2.8e-90;
Matches 585; Conservative 1; Mismatches 293; Indels 1; Gaps 1;

```
Qy    189 gaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtca 248
Db     40 GAGGAAAGTCTGTTACTACTACGACGGGGATGTTGGAAATTACTATTATGGACAAGGCCA 99

Qy    249 tcccatgaagccccatcgcatccgcatgacccatgcctcctcgctcactacggtctcct 308
Db    100 CCCAATGAAGCCTCACCGAATCCGCATGACTCATAATTTGCTGCTCAACTATGGTCTCTA 159

Qy    309 tcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttcca 368
Db    160 CCGAAAAATGGAAATCTATCGCCCTCACAAAGCCAATGCTGAGGAGATGACCAAGTACCA 219

Qy    369 cgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagat 428
Db    220 CAGCGATGACTACATTAAATTCTTGCGCTCCATCCGTCCAGATAACATGTCGGAGTACAG 279

Qy    429 tcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattc 488
Db    280 CAAGCAGATGCAGAGATTCAACGTTGGTGAGGACTGTCCAGTATTCGATGGCCTGTTTGA 339

Qy    489 cttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcct 548
Db    340 GTTCTGTCAGTTGTCTACTGGTGGTTCTGTGGCAAGTGCTGTGAAACTTAATAAGCAGCA 399

Qy    549 ctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctc 608
Db    400 GACGGACATCGCTGTGAATTGGGCTGGGGGCCTGCACCATGCAAAGAAGTCCGAGGCATC 459

Qy    609 tggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatga 668
Db    460 TGGCTTCTGTTACGTCAATGATATCGTCTTGGCCATCCTGGAACTGCTAAAGTATCACCA 519

Qy    669 gcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttta 728
Db    520 GAGGGTGCTGTACATTGACATTGATATTCACCATGGTGACGGCGTGGAAGAGGCCTTCTA 579

Qy    729 tgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtac 788
```

```
Db    580 CACCACGGACCGGGTCATGACTGTGTCCTTTCATAAGTATGGAGAGTACTTCCCAGGAAC 639

Qy    789 aggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccact 848
Db    640 TGGGGACCTACGGGATATCGGGGCTGGCAAAGGCAAGTATTATGCTGTTAACTACCCGCT 699

Qy    849 ggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagt 908
Db    700 CCGAGACGGGATTGATGACGAGTCCTATGAGGCCATTTTCAAGCCGGTCATGTCCAAAGT 759

Qy    909 tatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctgg 968
Db    760 AATGGAGATGTTCCAGCCTAGTGCGGTGGTCTTACAGTGTGGCTCAGACTCCCTATCTGG 819

Qy    969 ggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgt-caaattta 1027
Db    820 GGRTCGGTTAGGTTGCTTCAATCTAACTATCAAAGGACACGCCAAGTGTGTGGGAATTTG 879

Qy   1028 tgagatcgttcaatgttcccctactgctcttgggtggtgg 1067
Db    880 TCAAGAGCTTTAACCTGCCTATGCTGATGCTGGGAGGSGG 919
```

```
RESULT  13
AW396525
LOCUS       AW396525                656 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  sh28h04.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1016-6128 5' similar to SW:HDAC_ARATH O22446 HISTONE
            DEACETYLASE ;, mRNA sequence.
ACCESSION   AW396525
VERSION     AW396525.1  GI:6914995
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 656)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Seq primer: -40RP from Gibco
            High quality sequence stop: 407.
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1016-6128"
                     /clone_lib="Gm-c1016"
                     /tissue_type="immature flowers of field grown plants"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II XR; Site_1: EcoR1; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from immature flowers of field grown plants. The cDNA
                     library was prepared using the Stratagene pBluescript II
                     XR library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
```

(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."
BASE COUNT 164 a 124 c 159 g 207 t 2 others
ORIGIN

Query Match 21.8%; Score 393.8; DB 9; Length 656;
Best Local Similarity 75.0%; Pred. No. 4.5e-89;
Matches 491; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

```
Qy    152 gcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatg 211
          |||||||||  |  | || | |||||||||| |||||||| || || || ||||| ||||
Db      1 GTGGCAATTCTTTACCATCTGCACCTGATGGGGTGAAGAGAAAGGTGTGCTATTTTTATG 60

Qy    212 accctgaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatcc 271
          | || ||||| || |||||||| ||||| |||||||| || |||||||| |||||||| |
Db     61 ATCCGGAGGTTGGGAATTACTATTATGGGCAAGGTCACCCAATGAAGCCACATCGCATAC 120

Qy    272 gcatgacccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagc 331
          | ||||| ||||| || || || || || ||  | || || || ||||||||||  ||||
Db    121 GGATGACACATGCTCTTCTAGCCCATTATGGATTGCTACAACACATGCAGGTTCATAAGC 180

Qy    332 ccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttc 391
          | || |||||  | ||    || || || || || || || || || |||||  | || |
Db    181 CATTTCCTGCTAGAGACAAGGACCTTTGTCGATTTCATGCTGATGATTATGTTGCATTCC 240

Qy    392 tccgcagcattacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatg 451
          | || ||||| || |||||||| ||||| |||||  |  | ||||| || ||||| ||||
Db    241 TTCGAAGCATAACTCCTGAAACACAGCAGGATCATCTGAGGCAACTCAAACGCTTTAATG 300

Qy    452 ttggtgaagactgtcccgtctttgacggcctttattccttttgccagacctatgctggag 511
          |||||||||| ||||| || ||||| ||||| || || || || || || ||||| || |
Db    301 TTGGTGAAGATTGTCCTGTTTTTGATGGCCTCTACTCTTTCTGTCAAACATATGCAGGGG 360

Qy    512 gatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactggg 571
          | || ||||| ||  |||| ||| | || || | ||  || ||||||||  |||||||  |
Db    361 GTTCAGTTGGGGGTGCTGTAAAGTTAAATCATGACCAATGTGATATTGCTGTCAACTGNG 420

Qy    572 ctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgata 631
          |||||||  | ||||| || |||||||| ||||| ||||| || || || || |||||||
Db    421 CTGGTGGGTTGCATCATGCCAAGAAGTGTGAGGCTTCTGGATTTTGCTATGTTAATGATA 480

Qy    632 tcgtcttagctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattg 691
          | ||| | || ||| |||| || || |     ||||||||||||||| ||||| || ||||
Db    481 TAGTCCTTGCAATCTTAGAACTGCTCANAACACATGAGCGTGTTCTATATGTGGACATTG 540

Qy    692 atatccaccacggggatggagtggaggaggcattttatgctactgacagggttatgactg 751
          | || ||||| || || || |||||||| || || ||  |||||||||| || |||||||
Db    541 ACATTCACCATGGAGACGGTGTGGAGGAAGCTTTCTACACTACTGACAGAGTCATGACTG 600

Qy    752 tctcgtttcataaatttggtgattactttcccggtacaggtcacattcaggatat 806
          | |  |||| | |||||||||| || ||| | ||||| ||  |   ||  |||||
Db    601 TGTATTTTCCTTAATTTGGTGACTATTTTTCTGGTACTGGGGATGGTCGTGATAT 655
```

RESULT 14
AW308961
LOCUS AW308961 597 bp mRNA linear EST 02-DEC-2001
DEFINITION sf92d04.y1 Gm-c1019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1019-3584 5' similar to SW:HDAC_ARATH O22446 HISTONE DEACETYLASE ;, mRNA sequence.
ACCESSION AW308961
VERSION AW308961.1 GI:6724562
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 597)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 1527 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 446.
FEATURES Location/Qualifiers
source 1. .597
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1019-3584"
/clone_lib="Gm-c1019"
/tissue_type="immature seed coats of greenhouse grown plants"
/lab_host="DH10B (Gibco BRL)"
/note="Vector: pSPORT1 (Life Technologies); Site_1: Not I; Site_2: Sal I; This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mgs) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."
BASE COUNT 149 a 93 c 156 g 199 t
ORIGIN

Query Match 21.6%; Score 391.2; DB 9; Length 597;
Best Local Similarity 78.5%; Pred. No. 2e-88;
Matches 468; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

```
Qy    406 cctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaagactgt 465
          |||||||| ||||||||||  |  | || || ||| | || |||||||| ||||||||
Db      1 CCTGAAACGCAGCAAGATCAATTGAGACAGCTGAAGAGGTTTAATGTTGGCGAAGACTGC 60

Qy    466 cccgtctttgacggcctttattccttttgccagacctatgctggaggatctgttggtggc 525
          || || ||||| || ||||| || || |||||||| ||||| ||||| ||||||||||||
Db     61 CCTGTATTTGATGGTCTTTACTCTTTCTGCCAGACATATGCAGGAGGTTCTGTTGGTGGT 120

Qy    526 tctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggtctccat 585
          || | ||| | ||||| ||  | || |||||||| || || ||||||||||||| |||
Db    121 GCTCTAAAGTTGAACCATGGAGTATGTGATATTGCAATAAATTGGGCTGGTGGTCTACAT 180

Qy    586 cacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtcttagctatc 645
          || || |||||||| ||||| ||||| || || || || ||||| || ||  | |||||
Db    181 CATGCAAAGAAGTGTGAGGCTTCTGGGTTTTGCTATGTTAATGACATTGTGCTGGCTATT 240

Qy    646 ctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccaccacggg 705
```

```
Db    241 TTGGAACTTCTCAAAATACATGAGCGTGTTCTGTATGTGGACATTGATATCCACCATGGT 300

Qy    706 gatggagtggaggaggcattttatgctactgacagggttatgactgtctcgtttcataaa 765

Db    301 GATGGTGTAGAGGAGGCCTTTTACACCACCGATAGGGTCATGACTGTTTCGTTTCATAAG 360

Qy    766 tttggtgattactttcccggtacaggtcacattcaggatataggttatggtagcggaaag 825

Db    361 TTTGGGGATTACTTTCCTGGAACAGGTGATATTCGTGATATTGGATATGCTAAAGGGAAA 420

Qy    826 tactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcatctgtta 885

Db    421 TATTATTCACTAAATGTTCCCTTGGATGATGGAATTGATGATGAGAGCTATCAGTCCTTG 480

Qy    886 ttcaagcccatcatggggaaaagttatggaaatttccgaccaggggctgtggtattgcaa 945

Db    481 TTTAAGCCAATAATGGGAAAGGTTATGGAGATTTTTAGGCCCCGTGCTGTTGTATTACAA 540

Qy    946 tgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaa 1001

Db    541 TGTGGTGCTGACTCTTTATCTGGGGACAGGTTAGGTTGCTTCATCTTTCCATTAAA 596


RESULT  15
AL558916
LOCUS       AL558916                847 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL558916 LTI_NFL008_TC2 Homo sapiens cDNA clone CS0DJ007YP20 5
            prime, mRNA sequence.
ACCESSION   AL558916
VERSION     AL558916.1  GI:12903904
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 847)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .847
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ007YP20"
                     /clone_lib="LTI_NFL008_TC2"
                     /sex="male"
                     /tissue_type="T cells from T cell leukemia"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime  end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      217 a    195 c    224 g    207 t      4 others
ORIGIN


  Query Match           21.4%;  Score 386.4;  DB 9;  Length 847;
  Best Local Similarity   67.1%;  Pred. No. 3.8e-87;
  Matches  546;  Conservative    1;  Mismatches  267;  Indels    0;  Gaps    0;

Qy    189 gaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtca 248

Db     34 GAGGAAAGTCTGTTACTACTACGACGGGGATGTTGGAAATTACTATTATGGACAAGGCCA 93


Qy    249 tcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcct 308

Db     94 CCCAATGAAGCCTCACCGAATCCGCATGACTCATAATTTGCTGCTCAACTATGGTCTCTA 153

Qy    309 tcagcatatgcaggttctcaagccctccctgcccgcgaacgtgatctctgccgcttcca 368

Db    154 CCGAAAAATGGAAATCTATCGCCCTCACAAAGCCAATGCTGAGGAGATGACCAAGTACCA 213

Qy    369 cgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagat 428

Db    214 CAGCGATGACTACATTAAATTCTTGCGCTCCATCCGTCCAGATAACATGTCGGAGTACAG 273

Qy    429 tcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattc 488

Db    274 CAAGCAGATGCAGAGATTCAACGTTGGTGAGGACTGTCCAGTATTCGATGGCCTGTTTGA 333

Qy    489 cttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcct 548

Db    334 GTTCTGTCAGTTGTCTACTGGTGGTTCTGTGGCAAGTGCTGTGAAACTTAATAAGCAGCA 393

Qy    549 ctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctc 608

Db    394 GACGGACATCGCTGTGAATTGGGCTGGGGGCCTGCACCATGCAAAGAAGTCCGAGGCATC 453

Qy    609 tggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatga 668

Db    454 TGGCTTCTGTTACGTCAATGATATCGTCTTGGCCATCCTGGAACTGCTAAAGTATCACCA 513

Qy    669 gcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttta 728

Db    514 GAGGGTGCTGTACATTGACATTGATATTCACCATGGTGACGGCGTGGAAGAGGCCTTCTA 573

Qy    729 tgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtac 788

Db    574 CACCACGGACCGGGTCATGACTGTGTCCTTTCATAAGTATGGAGAGTACTTCCCAGGAAC 633

Qy    789 aggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccact 848

Db    634 TGGGGACCTACGGGATATCGGGGCTGGCAAAGGCAAGTATTATGCTGTTAACTACCCSCT 693

Qy    849 ggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagt 908

Db    694 CCSAGACGGGATTGATGACGAGTCCTATGAGGCCATTTTCAAGCCGGTCATGTCCAAAGT 753

Qy    909 tatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctgg 968

Db    754 AATGGAGATGTTCCASCCTAGTGCGGTGGTCTTACAGTGTGGCTCAGACTCCCTATCTGG 813

Qy    969 ggatcggttaggttgcttcaatctttcaatcaaa 1002

Db    814 GGRTCGGTTAGGTTGCTTCAATCTAACTATCAAA 847


Search completed: April 28, 2002, 18:51:38
Job time: 7628 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 28, 2002, 17:07:40 ; Search time 52.54 Seconds
(without alignments)
8448.035 Million cell updates/sec

Title: US-09-645-337-1
Perfect score: 1807
Sequence: 1 agagagcagctcccttcccc.....atgatgatatgatatgacaa 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	740.4	41.0	1943	4	US-09-282-305-7	Sequence 7, Appli
	2	698.6	38.7	2019	4	US-09-282-305-5	Sequence 5, Appli
	3	559.4	31.0	1826	4	US-09-282-305-1	Sequence 1, Appli
	4	498.2	27.6	2111	1	US-08-528-255A-2	Sequence 2, Appli
	5	498.2	27.6	2111	1	US-08-717-365-2	Sequence 2, Appli
	6	468	25.9	1576	4	US-09-282-305-9	Sequence 9, Appli
	7	384.2	21.3	1475	4	US-09-282-305-3	Sequence 3, Appli
	8	202.6	11.2	841	4	US-08-998-416-317	Sequence 317, App
c	9	85.8	4.7	7218	1	US-08-232-463-14	Sequence 14, Appl
	10	50.8	2.8	7218	1	US-08-232-463-14	Sequence 14, Appl
	11	44.6	2.5	289	4	US-09-007-005-17	Sequence 17, Appl
	12	44.6	2.5	289	4	US-09-244-796-17	Sequence 17, Appl
c	13	38.8	2.1	289	4	US-09-007-005-17	Sequence 17, Appl
c	14	38.8	2.1	289	4	US-09-244-796-17	Sequence 17, Appl
	15	38.8	2.1	485	4	US-09-177-325-4	Sequence 4, Appli
	16	38.8	2.1	485	4	US-09-411-812A-4	Sequence 4, Appli
	17	38.8	2.1	485	4	US-09-590-113-4	Sequence 4, Appli
	18	38.2	2.1	3095	6	5231168-1	Patent No. 5231168
	19	38	2.1	2518	4	US-09-433-699-3	Sequence 3, Appli
	20	37.8	2.1	966	2	US-08-766-738-2	Sequence 2, Appli
	21	37.8	2.1	3211	2	US-08-574-959A-8	Sequence 8, Appli
	22	37.8	2.1	3211	4	US-09-357-014-8	Sequence 8, Appli
	23	37.8	2.1	3901	2	US-08-574-959A-6	Sequence 6, Appli
	24	37.8	2.1	3901	4	US-09-357-014-6	Sequence 6, Appli
	25	37.4	2.1	696	4	US-09-461-697-193	Sequence 193, App
	26	37.4	2.1	699	4	US-09-461-697-191	Sequence 191, App
	27	37.4	2.1	717	4	US-09-461-697-189	Sequence 189, App
	28	37.4	2.1	774	4	US-09-461-697-187	Sequence 187, App
	29	37.4	2.1	819	4	US-09-461-697-185	Sequence 185, App
	30	37.4	2.1	1669	4	US-09-461-697-184	Sequence 184, App
c	31	37.4	2.1	53526	3	US-08-658-136-2	Sequence 2, Appli
c	32	37.4	2.1	53577	3	US-08-658-136-1	Sequence 1, Appli
	33	37.2	2.1	1924	3	US-08-961-083-159	Sequence 159, App
	34	37	2.0	2295	1	US-08-375-300-3	Sequence 3, Appli
	35	37	2.0	2295	3	US-09-177-431-3	Sequence 3, Appli
	36	37	2.0	2295	5	PCT-US95-16930-3	Sequence 3, Appli
	37	37	2.0	4080	1	US-08-375-300-1	Sequence 1, Appli
	38	37	2.0	4080	3	US-09-177-431-1	Sequence 1, Appli
	39	37	2.0	4080	5	PCT-US95-16930-1	Sequence 1, Appli
	40	36.4	2.0	282	4	US-09-461-697-205	Sequence 205, App
	41	36.4	2.0	306	4	US-09-461-697-203	Sequence 203, App
	42	36.4	2.0	2255	2	US-08-741-134-1	Sequence 1, Appli
	43	36.4	2.0	9636	1	US-08-323-170B-1	Sequence 1, Appli
	44	36.4	2.0	9636	4	US-08-954-441-1	Sequence 1, Appli
c	45	35.8	2.0	1505	1	US-07-915-246-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-282-305-7
; Sequence 7, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)..(1610)
US-09-282-305-7

Query Match 41.0%; Score 740.4; DB 4; Length 1943;
Best Local Similarity 75.0%; Pred. No. 3e-230;
Matches 953; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

```
Qy    144 ggatactggcggcaattcgctggcgtc---cggacctgatggtgtgaagaggaaagtttg 200
          ||   | ||||||||| || ||  ||||   ||| || || ||   |||| ||   || ||
Db     74 gggctccggcggcaactccctcccgtccgtcggccccgacgggcagaagcggcgcgtgtg 133

Qy    201 ttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaagcc 260
          || ||||| ||||| || || ||||| |||||||| || || || ||||| ||||||||
Db    134 ctacttctacgacccggatgtgggcaactactactacgggcagggccatccgatgaagcc 193

Qy    261 ccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatatgca 320
          || |||||||| ||||| ||  | || || || | | |||||| ||||  | || |||||
Db    194 gcaccgcatccggatgacgcactcgctgctggcgcgctacggcctcctcaaccagatgca 253

Qy    321 ggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgacta 380
          |||   |   |||  ||| |||||||| || || ||||||||||||||||||||||||
Db    254 ggtgtaccgccccaacccggcccgcgaccgcgacctctgccgcttccacgccgacgacta 313

Qy    381 tgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaacttaa 440
            ||  || || |||  | | || || ||||| ||||| || ||||| ||||  || ||
Db    314 catcaacttcctgcgctccgtcacgccggaaacgcagcaggaccagatccgcctgctcaa 373
```

```
Qy    441 gcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagac 500
          ||||||||| || || || ||||| |||||||| |||||||| ||   ||| ||||||||
Db    374 gcgcttcaacgtcggcgaggactgccccgtcttcgacggcctctacagcttctgccagac 433

Qy    501 ctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgc 560
          |||||| || |  || || || ||  | |||||||| ||||||||||     || || ||
Db    434 ctatgcgggcgcctccgtcggcgggccgtcaagctcaaccacggcc---atgacatcgc 490

Qy    561 catcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgtta 620
          |||||||| | || || || || ||||| ||||||||||||| || |||||||| ||
Db    491 aatcaactggtcgggggcctgcaccacgccaagaagtgcgaggcgtcgggcttctgcta 550

Qy    621 cgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttcttta 680
          ||||||||| |||||  | || || || ||||| || ||||| || ||| | ||||| ||
Db    551 cgtcaatgacatcgtgctcgccatactcgagctgctcaagcatcacgagagagttctgta 610

Qy    681 tgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactgacag 740
          ||||||||| |||||||||| || |||||||||||||||| || ||  | || || ||
Db    611 tgtcgatatcgatatccaccatggtgatggagtggaggaggctttctacacaacagatag 670

Qy    741 ggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattca 800
          |||||||||||||||||| || || |||||||||| || || || |||||  | || |
Db    671 ggttatgactgtctcgttccacaagtttggtgattatttcccaggaacaggggatatccg 730

Qy    801 ggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaat 860
          || || ||  |   |  || |||||||| || || ||||| || || |||||||| ||
Db    731 tgacattgggcactcaaaagggaagtactactccctgaatgtcctctagatgatgggat 790

Qy    861 cgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaatttt 920
          |||||||| ||||| ||     | || ||||| |||||||| |||||||||||  ||||
Db    791 tgatgatgaaagctaccagtccctttttaagccaatcatgggcaaagttatggaggtttt 850

Qy    921 ccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttagg 980
          ||| || || || || ||  | || |||||||||| ||| | ||||||||| |||| ||
Db    851 ccgccctggtgcagttgtgcttcagtgtggtgctgattccttgtctggggataggttggg 910

Qy    981 ttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaa 1040
          |||||||| || |||||||||||||||| || || || |  | |||||| || |||||
Db    911 ctgcttcaacctctcaatcaaaggtcatgcggaatgtgttaggtatatgaggtctttcaa 970

Qy   1041 tgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctg 1100
          |||||  |  ||||| | |||||||||||| || || ||  | |||||||| || || ||
Db    971 cgttccattgttgctccttggtggtggtggatataccataagaaatgttgcacgctgttg 1030

Qy   1101 gtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatgga 1160
          ||| || |||||||||||||| |||||    ||   |||||||||||||| |   ||||
Db   1031 gtgttatgagactggagttgctcttggccaagagcctgaagacaagatgcctgttaatga 1090

Qy   1161 atattatgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaaa 1220
          || |||||||||||| |||||||| || || ||||| ||||| ||||||||||||||| ||
Db   1091 gtactatgaatacttcggtccagattacactcttcatgttgcaccaagtaacatggagaa 1150

Qy   1221 taagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaa 1280
          || ||| | || ||    || || || || ||   | | |||||  | |||||| || ||
Db   1151 caaaaatacacgacaacaactggatgatatacgatctaaacttctggataatctttcaaa 1210

Qy   1281 gcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcccga 1340
          ||||  |||||||| ||||| |  ||||| || |||   ||||| ||||| |  || ||
Db   1211 acttcgacatgctcctagtgtccactttcaagagagagttcctgacacagaaatacctga 1270

Qy   1341 ggttgatgaagaccaagaagatggggataaaagatgggatccggattcagacatggatgt 1400
          |   |||||||| ||||| |||   ||| || ||   ||||| || || || ||||| ||
Db   1271 gcaagatgaagatcaagatgatccagatgaacgacatgatcctgactctgatatggaagt 1330

Qy   1401 tgatgatgac 1410
          |||||  ||
Db   1331 ggatgaccac 1340


RESULT   2
US-09-282-305-5
; Sequence 5, Application US/09282305
; Patent No. 6287843
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 2019
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (140)..(1459)
US-09-282-305-5


  Query Match             38.7%;  Score 698.6;  DB 4;  Length 2019;
  Best Local Similarity   73.0%;  Pred. No. 1.2e-216;
  Matches  912;  Conservative    0;  Mismatches  334;  Indels    3;  Gaps    1;

Qy    162 gctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgaggt 221
          || | |  | ||  | || ||   |||| |    || || || ||||| ||| | |||||
Db    175 gccgaccactggcgcggacgggtcgaagcgccgcgtctgctacttctacgacgcggaggt 234

Qy    222 cggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgaccca 281
          |||||  |||||||| || || || || || |||||||| || ||||||||||||||||
Db    235 gggcaactactactacgggcagggccacccgatgaagccgcaccgcatccgcatgaccca 294

Qy    282 tgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagcccttccctgc 341
          || || |||| | |||||| |||||  | || ||||| ||  ||   ||   ||||||
Db    295 cgcgctgctcggccgctacggcctcctcgaccagatgcaagtgttccgccctcaccctgc 354

Qy    342 ccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagcat 401
          |||||| || || |||||||||||||||||||||||| || ||||| || |||||   | |
Db    355 ccgcgaccgcgacctctgccgcttccacgccgacgattacgtctccttcctccggtccgt 414

Qy    402 tacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaaga 461
          |||||  |||||  |||||  || |||||  |||   || ||||||||||| || || || ||
Db    415 cacccccgaaacgcagcaggaccagatccgcgcgctcaagcgcttcaacgtcggcgagga 474

Qy    462 ctgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgttgg 521
          |||  |||||||| |||||  || ||    || || ||||| || || || || ||||||||
Db    475 ctgccccgtcttcgacggtctctacagtttctgtcagacgtacgcgggggctctgttgg 534

Qy    522 tggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggtct 581
          |||  | |||||||  |||||  ||||     |||||  ||||||||||||||| || || ||
Db    535 cggcgccgtcaagctcaaccatggccat---gatatcgccatcaactgggccggcggact 591

Qy    582 ccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtcttagc 641
          |||  |||||  |||||||||||||||||  || || || || || |||||  || ||| | ||
Db    592 ccaccacgccaagaagtgcgaggcctccgggttttgctatgttaatgacattgtcctcgc 651

Qy    642 tatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccacca 701
          |||||  |||||||| ||| | ||  |||| |||||  || || || ||||||||||||||
Db    652 catcctcgagctcctcaagtaccaccagcgcgttctgtacgtggacattgatatccacca 711

Qy    702 cggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgtttca 761
          ||||||  || ||||||||||| |||||| | || ||| |||| ||||| ||||| || ||
Db    712 cggggacggcgtggaggaggcttttataccacagaccgggtgatgacagtctcattcca 771

Qy    762 taaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagcgg 821
          || ||||| ||||| || || || |||||  ||||||  ||| | ||  |    |  ||
Db    772 caagtttggagattatttccctgggacaggggacattcgtgatgttgggcactcaaaggg 831

Qy    822 aaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcatct 881
          || || || || || ||||| || ||||| ||||| || ||||||||||||||| ||
Db    832 taaatattactccctgaatgttcccctggacgatggtattgatgatgagagctaccagtc 891
```

```
Qy    882 gttattcaagcccatcatggggaaagttatggaaatttccgaccaggggctgtggtatt 941
Db    892 gttgttcaagccaataatgggcaaggtgatggaggtcttcaaccctggtgcagtcgtgct 951

Qy    942 gcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaa 1001
Db    952 ccagtgtggtgcggattcattgtcgggtgacaggttgggctgtttcaacctctctattaa 1011

Qy   1002 aggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttggg 1061
Db   1012 gggtcacgcagaatgtgtaagatttatgaggtccttcaacgtcccgctgttgctgcttgg 1071

Qy   1062 tggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagttgc 1121
Db   1072 tggtggtgggtataccataagaaacgttgcacggtgttggtgctacgagacaggagttgc 1131

Qy   1122 acttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggtcc 1181
Db   1132 ccttggtcatgagctcactgacaagatgccacctaatgagtactatgagtattttggtcc 1191

Qy   1182 agactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatgct 1241
Db   1192 agattacactctacatgtcgctccaagtaacatggagaataaaaacacacggcatcaatt 1251

Qy   1242 tgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagtgt 1301
Db   1252 ggatgacataaaatcaaaacttctagataatctttcaaaactccgacatgctcctagtgt 1311

Qy   1302 accatttcaggaaagaccacctgatacagagactcccgaggttgatgaagaccaagaaga 1361
Db   1312 tcagtttcaagagcgacctcctgaggctgagttacctgagcaagatgaagacaaagagaa 1371

Qy   1362 tggggataaaagatgggatccggattcagacatggatgttgatgatgac 1410
Db   1372 tcctgatgaaagacatgatgctgattctgatgtggagatgaatgatgcc 1420
```

RESULT 3
US-09-282-305-1
; Sequence 1, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1826
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (29)..(1405)
US-09-282-305-1

Query Match 31.0%; Score 559.4; DB 4; Length 1826;
Best Local Similarity 66.8%; Pred. No. 2.1e-171;
Matches 797; Conservative 0; Mismatches 396; Indels 0; Gaps 0;

```
Qy    160 tcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgag 219
Db     62 tctccggcgggcggggaggatgcgcaccgccgccgcgtcagctatttctacgagccgtcg 121

Qy    220 gtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgacc 279
Db    122 atcggagactactactacgggcaaggtcacccgatgaagccccaccgcatccgaatggcg 181

Qy    280 catgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagccctccct 339
Db    182 cactcgctggtggtccactacggcctccaccgcctcctcgagctctcccgccctacccg 241

Qy    340 gcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagc 399
Db    242 gcctctgaggccgacatccgccgcttccactccgacgactacgtcgctttcctcgcgtcc 301

Qy    400 attacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaa 459
Db    302 gccaccggaaacccgggtgtcctcgacccgcgcgccattaagcgctttaacgtcggtgag 361

Qy    460 gactgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgtt 519
Db    362 gactgccccgtgttcgacggtctcttcccctctgccaggcctccgctgggggaagcatc 421

Qy    520 ggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggt 579
Db    422 ggcgccgccgtcaagcttaaccgcggggacgccgacatcaccgtcaactgggcgggcggc 481

Qy    580 ctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtctta 639
Db    482 ctccaccacgccaagaagagcgaggcctccgggttctgctacgtcaacgacatcgtcctc 541

Qy    640 gctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccac 699
Db    542 gccatcctcgagctcctcaagttccacaggcgtgtgctatatgtagacattgatgtccac 601

Qy    700 cacggggatggagtggaggaggcatttttatgctactgacagggttatgactgtctcgttt 759
Db    602 catggagatggcgtggaggaggccttcttcactacaaaccgagtcatgactgtttccttt 661

Qy    760 cataaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagc 819
Db    662 cacaagtatggggatttttccctggtactggacatatcactgacgttggggcagccgaa 721

Qy    820 ggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcat 879
Db    722 gggaagcattatgctctgaatgttcccctgagtgatggtatcgatgacaccacctttcgt 781

Qy    880 ctgttattcaagcccatcatggggaaagttatggaaattttccgaccaggggctgtggta 939
Db    782 ggtctgtttcaatgcatcattaagaaagttatggaggtttatcagccagacgtggttgtc 841

Qy    940 ttgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatc 999
Db    842 ctccaatgcggagctgactctttggctggagacaggttaggttgcttcaacctgtctgtg 901

Qy   1000 aaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttg 1059
Db    902 aagggtcatgctgactgcctccgtttccttaggtcgtacaatgttcctatgatggtttta 961

Qy   1060 ggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagtt 1119
Db    962 ggtggtggaggttacaccatcagaaatgttgcacgctgctggtgctacgagaccgcagtt 1021

Qy   1120 gcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggt 1179
Db   1022 gctgttggagttgaacctgataacaagctgccttacaatgattactatgagtactttggc 1081

Qy   1180 ccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatg 1239
Db   1082 cctgattatactcttcatatccaaccaaaaagtgttgaaaacctgaataccacaaaggac 1141

Qy   1240 cttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagt 1299
Db   1142 ttggagaacataaagaacatgatattggagaacctgtcaaagatagaacatgttcccagc 1201

Qy   1300 gtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaaga 1352
```

```
Db   1202 actcaattccatgacagaccgtcagaccctgaagctccagaggagaaagagga 1254


RESULT   4
US-08-528-255A-2
; Sequence 2, Application US/08528255A
; Patent No. 5659016
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  :  FURUKAWA, YOICHI
;    TITLE OF INVENTION:  RPDL PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/528,255A
;      FILING DATE:  September 14, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP6-227876
;      FILING DATE:  22-SEPTEMBER-1994
;      APPLICATION NUMBER:  JP7-183763
;      FILING DATE:  20-JULY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1335
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2111
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  human fetal lung cDNA library
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  64..1512
;      IDENTIFICATION METHOD:  experimental examination
US-08-528-255A-2


  Query Match           27.6%;  Score 498.2;  DB 1;  Length 2111;
  Best Local Similarity   64.0%;  Pred. No. 1.8e-151;
  Matches  752;  Conservative    0;  Mismatches  423;  Indels    0;  Gaps    0;

Qy    189 gaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtca 248
          ||||||||| ||||| | ||| |||   || || || |||||||| ||||| ||||| ||
Db     87 GAGGAAAGTCTGTTACTACTACGACGGGGATGTTGGAAATTACTATTATGGACAAGGCCA 146

Qy    249 tcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcct 308
          || |||||||| || || |||||||||||| |||    | ||     |||| ||||||
Db    147 CCCAATGAAGCCTCACCGAATCCGCATGACTCATAATTTGCTGCTCAACTATGGTCTCTA 206

Qy    309 tcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttcca 368
          |   | ||| |  |       ||   |   |||   |     ||  |   |   | |||

Db    207 CCGAAAAATGGAAATCTATCGCCCTCACAAAGCCAATGCTGAGGAGATGACCAAGTACCA 266

Qy    369 cgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagat 428
          |  ||| |||||  |    ||  | |||  |||    || || | |  |   ||  | |
Db    267 CAGCGATGACTACATTAAATTCTTGCGCTCCATCCGTCCAGATAACATGTCGGAGTACAG 326

Qy    429 tcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattc 488
          ||  |  || | ||||| ||||||| |||||||| || || || ||||| | |
Db    327 CAAGCAGATGCAGAGATTCAACGTTGGTGAGGACTGTCCAGTATTCGATGGCCTGTTTGA 386

Qy    489 ctttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcct 548
          || || |||   | | |||| || ||||| |   |  |||| || |||||  |    |
Db    387 GTTCTGTCAGTTGTCTACTGGTGGTTCTGTGGCAAGTGCTGTGAAACTTAATAAGCAGCA 446

Qy    549 ctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctc 608
          || || ||  | || |||||||| || || || || || ||||||| |||||| ||
Db    447 GACGGACATCGCTGTGAATTGGGCTGGGGGCCTGCACCATGCAAAGAAGTCCGAGGCATC 506

Qy    609 tggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatga 668
          ||||||||||||||||||||||||||||||| || ||||| || || || ||| | ||  |
Db    507 TGGCTTCTGTTACGTCAATGATATCGTCTTGGCCATCCTGGAACTGCTAAAGTATCACCA 566

Qy    669 gcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttta 728
          | | || || ||  | || |||||||| ||||| || || || ||||| ||||| || ||
Db    567 GAGGGTGCTGTACATTGACATTGATATTCACCATGGTGACGGCGTGGAAGAGGCCTTCTA 626

Qy    729 tgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtac 788
          | || ||| |||| |||||||| || |||||||| | ||| || ||||| || || ||
Db    627 CACCACGGACCGGGTCATGACTGTGTCCTTTCATAAGTATGGAGAGTACTTCCCAGGAAC 686

Qy    789 aggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccact 848
          ||  || | | |||||| ||   ||| |  || ||||| ||| || | ||    || ||
Db    687 TGGGGACCTACGGGATATCGGGGCTGGCAAAGGCAAGTATTATGCTGTTAACTACCCGCT 746

Qy    849 ggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagt 908
          || || || ||||| ||| |||| |     | ||||||||  |||||   |||||
Db    747 CCGAGACGGGATTGATGACGAGTCCTATGAGGCCATTTTCAAGCCGGTCATGTCCAAAGT 806

Qy    909 tatggaaatttttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctgg 968
          ||||| || |||  ||  | || ||||| || || ||||  | ||||||||||||||
Db    807 AATGGAGATGTTCCAGCCTAGTGCGGTGGTCTTACAGTGTGGCTCAGACTCCCTATCTGG 866

Qy    969 ggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttat 1028
          ||||||||||||||||||||||  | |||||||| || ||  |||| ||  ||||| |
Db    867 GGATCGGTTAGGTTGCTTCAATCTAACTATCAAAGGACACGCCAAGTGTGTGGAAATTTGT 926

Qy   1029 gagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaatgt 1088
          |     || ||  | ||  | ||| |  |||| || ||||||||||| || || || ||
Db    927 CAAGAGCTTTAACCTGCCTATGCTGATGCTGGGAGGCGGTGGTTACACCATTCGTAACGT 986

Qy   1089 tgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaagacaagat 1148
          |||||| ||| ||   || ||||| |  || || || |     ||  |     |  || |
Db    987 TGCCCGGTGCAGGACATATGAGACAGCTGTGGCCCTGGATACGGAGATCCCTAATGAGCT 1046

Qy   1149 gccggagcatgaatattatgaatactttggtccagactatacacttcacgttgctccaag 1208
          ||  |  |||| || | |||||||||||| |||| |  |  || ||| |   |||
Db   1047 TCCATACAATGACTACTTTGAATACTTTGGACCAGATTTCAAGCTCCACATCAGTCCTTC 1106

Qy   1209 taacatggaaaataagaattctcgtcagatgcttgaagagattcgcaatgaccttctcca 1268
          || |||   ||  ||||  |   | ||   || ||  ||||     |   ||  |  |
Db   1107 CAATATGACTAACCAGAACACGAATGAGTACCTGGAGAAGATCAAACAGCGACTGTTTGA 1166

Qy   1269 caatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatac 1328
          || ||    | ||| | ||| || ||  | || | | | ||||  |  ||    ||  |
Db   1167 GAACCTTAGAATGCTGCCGCACGCACCTGGGGTCCAAATGCAGGCGATTCCTGAGGACGC 1226

Qy   1329 agagactcccgaggttgatgaagaccaagaagatg 1363
             ||   |||  ||  || ||  | ||||| |
Db   1227 CATCCCTGAGGAGAGTGGCGATGAGGACGAAGACG 1261


RESULT   5
US-08-717-365-2
```

```
; Sequence 2, Application US/08717365
; Patent No. 5763182
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  :  FURUKAWA, YOICHI
;    TITLE OF INVENTION:  RPDL PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,365
;      FILING DATE:  23-SEP-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/528 255
;      FILING DATE:  September 14, 1995
;      APPLICATION NUMBER:  JP6-227876
;      FILING DATE:  22-SEPTEMBER-1994
;      APPLICATION NUMBER:  JP7-183763
;      FILING DATE:  20-JULY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1335
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2111
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  human fetal lung cDNA library
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  64..1512
;      IDENTIFICATION METHOD:  experimental examination
US-08-717-365-2


  Query Match           27.6%;  Score 498.2;  DB 1;  Length 2111;
  Best Local Similarity  64.0%;  Pred. No. 1.8e-151;
  Matches  752;  Conservative    0;  Mismatches  423;  Indels    0;  Gaps    0;

Qy     189 gaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtca 248
           ||||||||| ||||| | ||| |||   || || || |||||||| ||||| ||||| ||
Db      87 GAGGAAAGTCTGTTACTACTACGACGGGGATGTTGGAAATTACTATTATGGACAAGGCCA 146

Qy     249 tcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcct 308
           || |||||||| || || |||||||||||| |||    | ||     |||| ||||||
Db     147 CCCAATGAAGCCTCACCGAATCCGCATGACTCATAATTTGCTGCTCAACTATGGTCTCTA 206

Qy     309 tcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttcca 368
           |   | ||| |  |       ||   |  |||   |    ||  |   |   | |||
Db     207 CCGAAAAATGGAAATCTATCGCCCTCACAAAGCCAATGCTGAGGAGATGACCAAGTACCA 266

Qy     369 cgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagat 428
           |  ||| |||||  |    ||  | |||  |||    || || | |  |   ||  | |
Db     267 CAGCGATGACTACATTAAATTCTTGCGCTCCATCCGTCCAGATAACATGTCGGAGTACAG 326

Qy     429 tcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacqqcctttattc 488
           ||  |  || | ||||| |||||||| |||||||| || || || ||||| | |
Db     327 CAAGCAGATGCAGAGATTCAACGTTGGTGAGGACTGTCCAGTATTCGATGGCCTGTTTGA 386

Qy     489 cttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcct 548
           || || |||   | | |||| || ||||| |   |  |||| || ||||| |   |
Db     387 GTTCTGTCAGTTGTCTACTGGTGGTTCTGTGGCAAGTGCTGTGAAACTTAATAAGCAGCA 446

Qy     549 ctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctc 608
           || || ||  | || |||||||| || || || || || |||||| ||||||
Db     447 GACGGACATCGCTGTGAATTGGGCTGGGGGCCTGCACCATGCAAAGAAGTCCGAGGCATC 506

Qy     609 tggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatga 668
           |||||||||||||||||||||||||||||||| || ||||| || || || ||| | ||  |
Db     507 TGGCTTCTGTTACGTCAATGATATCGTCTTGGCCATCCTGGAACTGCTAAAGTATCACCA 566

Qy     669 gcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttta 728
           | | || || ||  | || |||||||| ||||| || || || ||||| ||||| || ||
Db     567 GAGGGTGCTGTACATTGACATTGATATTCACCATGGTGACGGCGTGGAAGAGGCCTTCTA 626

Qy     729 tgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtac 788
           | || ||| |||| |||||||| || |||||||| | ||| || ||||| || || ||
Db     627 CACCACGGACCGGGTCATGACTGTGTCCTTTCATAAGTATGGAGAGTACTTCCCAGGAAC 686

Qy     789 aggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccact 848
           ||  || | | |||||| ||   ||| | || ||||| ||| || | ||    || ||
Db     687 TGGGGACCTACGGGATATCGGGGCTGGCAAAGGCAAGTATTATGCTGTTAACTACCCGCT 746

Qy     849 ggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagt 908
           || || || ||||| |||  |||| |      | ||||||||  |||||  |||||
Db     747 CCGAGACGGGATTGATGACGAGTCCTATGAGGCCATTTTCAAGCCGGTCATGTCCAAAGT 806

Qy     909 tatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctgg 968
           ||||| || |||| ||  | || ||||| || || |||||  | ||||||||||||||
Db     807 AATGGAGATGTTCCAGCCTAGTGCGGTGGTCTTACAGTGTGGCTCAGACTCCCTATCTGG 866

Qy     969 ggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttat 1028
           |||||||||||||||||||||  | |||||||| || ||  |||| ||  ||||| |
Db     867 GGATCGGTTAGGTTGCTTCAATCTAACTATCAAAGGACACGCCAAGTGTGTGGAATTTGT 926

Qy    1029 gagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaatgt 1088
           |      || ||  | ||  | ||| |  |||| || ||||||||||| || || || ||
Db     927 CAAGAGCTTTAACCTGCCTATGCTGATGCTGGGAGGCGGTGGTTACACCATTCGTAACGT 986

Qy    1089 tgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaagacaagat 1148
           |||||| ||| ||   || ||||| |  || || || |     ||  |     |  || |
Db     987 TGCCCGGTGCAGGACATATGAGACAGCTGTGGCCCTGGATACGGAGATCCCTAATGAGCT 1046

Qy    1149 gccggagcatgaatattatgaatactttggtccagactatacacttcacgttgctccaag 1208
           ||  |  |||| || | ||||||||||| ||||| |  |  || ||| |   |||
Db    1047 TCCATACAATGACTACTTTGAATACTTTGGACCAGATTTCAAGCTCCACATCAGTCCTTC 1106

Qy    1209 taacatggaaaataagaattctcgtcagatgcttgaagagattcgcaatgaccttctcca 1268
           || |||   ||  |||| |  | ||  || ||  ||||     |   ||  |  |
Db    1107 CAATATGACTAACCAGAACACGAATGAGTACCTGGAGAAGATCAAACAGCGACTGTTTGA 1166

Qy    1269 caatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatac 1328
           || ||    | ||| | ||| || ||  | || | | | ||||  |  ||    ||  |
Db    1167 GAACCTTAGAATGCTGCCGCACGCACCTGGGGTCCAAATGCAGGCGATTCCTGAGGACGC 1226

Qy    1329 agagactcccgaggttgatgaagaccaagaagatg 1363
           ||   |||  ||  || ||  | ||||  |
Db    1227 CATCCCTGAGGAGAGTGGCGATGAGGACGAAGACG 1261


RESULT   6
US-09-282-305-9
; Sequence 9, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
```

```
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 1576
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (38)..(1336)
US-09-282-305-9


  Query Match           25.9%;  Score 468;  DB 4;  Length 1576;
  Best Local Similarity   63.1%;  Pred. No. 9.9e-142;
  Matches  739;  Conservative    0;  Mismatches  430;  Indels    3;  Gaps    1;

Qy    217 gaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatg 276
          || || ||||||  ||||| ||| | |  ||| ||||||||||| |||||  || | |||
Db     83 gatgttggcaatgtctactttgggccaaatcaccccatgaagccacatcgtctctgtatg 142

Qy    277 acccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagcccttc 336
          || |||  |||  |   |   || || || | |||  | ||| || |    | |||   |
Db    143 acacatcaccttgttctttcatatggacttcatcaaaagatggagatatataggccacac 202

Qy    337 cctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgc 396
            ||      |  || ||  ||  |||||  | |  || |||||    ||  | | |
Db    203 aaagcatatccaatagagcttgcccaattccattctgctgattatgtggaattcttgcac 262

Qy    397 agcattacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggt 456
           | || || |||||  |||||||     |     |  |||| |   | | |||| ||||
Db    263 cggataactcctgattcccagcacctatatgcaagtgaactaactagatacaatcttgga 322

Qy    457 gaagactgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatct 516
          ||||||||||| ||||||||     | | |   || ||||| | |||||| |||||| ||
Db    323 gaagactgtccggtctttgataatttgtttgagttctgccaaatctatgcgggaggaact 382

Qy    517 gttggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggt 576
           | | ||   ||  ||   | || ||     | || || |||||||| || |||||||||
Db    383 ttagatgctgctcgcagattaaatcataaaatatgtgacattgccattaattgggctggt 442

Qy    577 ggtctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtc 636
          || || ||||| || || ||||| ||||| || ||||||||||||| | |||||| | ||
Db    443 gggctacatcatgccaaaaagtgtgaggcttcaggcttctgttacattaatgatctagta 502

Qy    637 ttagctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatc 696
          ||||  || || ||||| || ||| | ||||    | |||||||| | || |||||| ||
Db    503 ttaggaattctggagcttctcaagtaccatgccagggttctttatattgacattgatgtc 562

Qy    697 caccacggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcg 756
          || || ||  |||||||| || || || ||||||   |||||||||| ||||||||
Db    563 catcatggagatggagttgaagaagccttttatttcactgacagggtaatgactgtgagt 622

Qy    757 tttcataaatttggtga---ttacttttcccggtacaggtcacattcaggatataggttat 813
          || || ||  | ||||||    | |||||| || |||||| | ||| ||||||||||  |
Db    623 ttccacaagtatggtgacctgttctttcctggaacaggtgatattaaggatataggagaa 682

Qy    814 ggtagcggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagc 873
           |    ||||| || ||| || ||||  | |||||  | ||||| || |||||    |||
Db    683 agggaaggaaaatattatgctatcaacattccacttaaagatgggatagatgacactagc 742

Qy    874 tatcatctgttattcaagcccatcatggggaaagttatggaaatttttccgaccaggggct 933
          | |  || |   || ||  | || || |  |||||| | || |  |  |  || || |||
Db    743 tttactcggccttttaaaacaattattgccaaagttgttgagacatatctgcctggtgct 802

Qy    934 gtggtattgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctt 993
           | ||  | |||||||| ||||| ||  |  |  ||||||| ||||| ||||||||||
Db    803 attgttcttcaatgtggggctgattcattggcgagggatcgtttaggctgcttcaatctc 862

Qy    994 tcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactg 1053
          || ||  |||| |||||||| || || || ||| | |     |||||| ||||||| |||
Db    863 tctattgaaggccatgctgaatgtgtaaagtttgtcaagaaattcaatattcccctctg 922

Qy   1054 ctcttgggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagact 1113
           |    || |||||||| ||||| |     ||||| || || || |||      || |||
Db    923 gtaactggaggtggtggatacaccaaggagaatgtagcacggtgttgggctgttgaaact 982

Qy   1114 ggagttgcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatac 1173
          || ||      | |     ||| |   | |  |||| ||  |  |||||||  ||| |||
Db    983 ggggtccttttagacacagaactcccaaatgagattccaaaaaatgaatatattgagtac 1042

Qy   1174 tttggtccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgt 1233
          |||| |||||| |||||| |  | ||| |       |||||||| |||  |||   |
Db   1043 tttgctccagattatacattgaaagttccaaatttgaacatggacaatttgaacagtaag 1102

Qy   1234 cagatgcttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgct 1293
             ||          ||          |  |  |  | | | |     |  | |||||||||
Db   1103 acctatctcagttcaatcaaagtgcaagtgatggagagtttgcggtacatacagcatgct 1162

Qy   1294 ccaagtgtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaagac 1353
          ||  |||| | | || ||    || || |||    | |  || ||  ||||||||||
Db   1163 cctggtgttcaaatgcaagaggttcctcccgatttttatatcccggactttgatgaagat 1222

Qy   1354 caagaagatggggataaaagatgggatccgga 1385
           ||   |||   ||| || |    || | | |
Db   1223 gaattggatcctgatgaacgtgttgaccagca 1254


RESULT   7
US-09-282-305-3
; Sequence 3, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (29)..(1084)
US-09-282-305-3


  Query Match           21.3%;  Score 384.2;  DB 4;  Length 1475;
  Best Local Similarity   68.6%;  Pred. No. 1.7e-114;
  Matches  530;  Conservative    0;  Mismatches  243;  Indels    0;  Gaps    0;

Qy    580 ctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtctta 639
           | ||| ||||| |||||| |||||||||| || ||||| |||||||| || |||||| |
Db    161 ccccaccacgccaagaagagcgaggcctccgggttctgctacgtcaacgacatcgtcctc 220

Qy    640 gctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccac 699
          || ||||| |||||||| |||   ||   |||||| || ||||| || |||||| ||||
Db    221 gccatcctcgagctcctcaagttccacaggcgtgtgctatatgtagacattgatgtccac 280

Qy    700 cacggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgttt 759
          || || ||||| |||||||||||| || |   ||||  || | || |||||||| || |||
```

```
Db    281 catggagatggcgtggaggaggccttcttcactacaaaccgagtcatgactgtttccttt 340

Qy    760 cataaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagc 819

Db    341 cacaagtatggggatttttttccctggtactggacatatcactgacgttggggcagccgaa 400

Qy    820 ggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcat 879

Db    401 gggaagcattatgctctgaatgttcccctgagtgatggtatcgatgacaccacctttcgt 460

Qy    880 ctgttattcaagcccatcatggggaaagttatggaaattttccgaccaggggctgtggta 939

Db    461 ggtctgtttcaatgcatcattaagaaagttatggaggtttatcagccagacgtggttgtc 520

Qy    940 ttgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatc 999

Db    521 ctccaatgcggagctgactctttggctggagacaggttaggttgcttcaacctgtctgtg 580

Qy   1000 aaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttg 1059

Db    581 aagggtcatgctgactgcctccgtttccttaggtcgtacaatgttcctatgatggtttta 640

Qy   1060 ggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagtt 1119

Db    641 ggtggtggaggttacaccatcagaaatgttgcacgctgctggtgctacgagaccgcagtt 700

Qy   1120 gcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggt 1179

Db    701 gctgttggagttgaacctgataacaagctgccttacaatgattactatgagtactttggc 760

Qy   1180 ccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatg 1239

Db    761 cctgattatactcttcatatccaaccaaaaagtgttgaaaacctgaataccacaaaggac 820

Qy   1240 cttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagt 1299

Db    821 ttggagaacataaagaacatgatattggagaacctgtcaaagatagaacatgttcccagc 880

Qy   1300 gtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaaga 1352

Db    881 actcaattccatgacagaccgtcagaccctgaagctccagaggagaaagagga 933


RESULT   8
US-08-998-416-317
; Sequence 317, Application US/08998416
; Patent No. 6239264
;  GENERAL INFORMATION:
;    APPLICANT:  Philippsen, Peter
;    APPLICANT:  Pohlmann, Rainer
;    APPLICANT:  Steiner, Sabine
;    APPLICANT:  Mohr, Christine
;    APPLICANT:  Wendland, Jurgen
;    APPLICANT:  Knechtle, Philipp
;    APPLICANT:  Rebischung, Corinne
;    TITLE OF INVENTION:  GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
;    TITLE OF INVENTION:  AND USES THEREOF
;    NUMBER OF SEQUENCES:  1152
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 6239264artis Corporation
;      STREET:  3054 Cornwallis Road
;      CITY:  Research Triangle Park
;      STATE:  No. 6239264th Carolina
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,416
;      FILING DATE:  24-DEC-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  CH 0016/97
;      FILING DATE:  31-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meigs, J. Timothy
;      REGISTRATION NUMBER:  38,241
;      REFERENCE/DOCKET NUMBER:  PF/5-30306/A/CGC1976
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-541-8587
;      TELEFAX:  919-541-8689
;  INFORMATION FOR SEQ ID NO:  317:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  841 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      ORGANISM:  PAG1256UP
US-08-998-416-317


  Query Match             11.2%;  Score 202.6;  DB 4;  Length 841;
  Best Local Similarity   62.7%;  Pred. No. 1.3e-55;
  Matches  349;  Conservative    0;  Mismatches  204;  Indels    4;  Gaps    2;

Qy    562 atcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgttac 621

Db      2 ATCAACTGGTCGGGCGGGCTGCACCACGCCAAGAAGAGCAATCCTTCGGGGTTCTGTTAC 61

Qy    622 gtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttctttat 681

Db     62 GTGAACGACATTGTTCTGGCGATTCTGAATCTGCTGCGCTACCACCCACGCGTTCTGTAC 121

Qy    682 gtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactgacagg 741

Db    122 ATTGACATTGATCTGCACCACGGAGACGGTGTCCAAGAAGCATTCTACACTACTGACCGC 181

Qy    742 gttatgactgtctcgtttcataaat---ttggtgattactttcccggtacaggtcacatt 798

Db    182 GTGTTCACGGTCTCGTTCCACAAGTACAATGGTGAGTTTTTTCCGGGAACGGGGGATTTG 241

Qy    799 caggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatgga 858

Db    242 GATGAGATCGGATGCTCGCGCGGCAAGCACTTTTCGCTGAATGTGCCGCTCAATGACGGC 301

Qy    859 atcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaatt 918

Db    302 ATCGATGATGATTCGTACATCAACTTATTTAAGAGCATCATAGACCCGCTAGTTACATCA 361

Qy    919 ttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggtta 978

Db    362 TACAAGCCAACAGTAATTATTCAGCAATGTGGAGCAGACTCTTTGGGGCATGACAGACTG 421

Qy    979 ggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttc 1038

Db    422 GGGTGTTTCAATCTAAATATCAGAGCCCACGGCGAGTGCGTC-AATTTGTGAAGTCGTTC 480

Qy   1039 aatgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgc 1098

Db    481 GGGATACCTATGCTATGTGTCGGTGGTGGAGGTTACACCCCCAGGAATGTGTCGCGGCTA 540

Qy   1099 tggtgctacgagactgg 1115

Db    541 TGGACGTACGAGACAGG 557


RESULT   9
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             4.7%;  Score 85.8;  DB 1;  Length 7218;
  Best Local Similarity   2.5%;  Pred. No. 5.8e-17;
  Matches    9;  Conservative  243;  Mismatches  115;  Indels    0;  Gaps    0;

Qy   1310 aggaaagaccacctgatacagagactcccgaggttgatgaagaccaagaagatggggata 1369
          :::::::::  :   :: : :::::     ::::  :: :::::  ::::::: ::::: :
Db   1413 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1354

Qy   1370 aaagatgggatccggattcagacatggatgttgatgatgaccgtaaacctataccaagca 1429
          ::::: ::::   :::   ::: : ::: :  :: :: ::  : :::   : :  ::: :
Db   1353 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1294

Qy   1430 gagtaaaaagagaagctgttgaaccagatacaaaggacaaggatggactgaaaggaatta 1489
          ::: ::::::::::::  :  :::  ::: : :::::: ::::: :::  :::::::: :
Db   1293 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1234

Qy   1490 tggagcgtggaaaaggttgtgaggtggaggtggatgagagtggaagcactaaggttacag 1549
           :::: : ::::::::  : :::: ::::: ::: ::::: ::::: :  ::::  : ::
Db   1233 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1174

Qy   1550 gagtaaacccagtgggagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaaca 1609
          ::: :::   :: ::::: :::::::: ::: : ::::: :::::::::::::: ::: :
Db   1173 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1114

Qy   1610 agggtggggcggagcaggcgtttcctcctaaaacataagactcggagcttctaatttctt 1669
          :::: :::: :::: ::: :          :::: : ::::   :::: |   ||  ||
Db   1113 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCC 1054

Qy   1670 gctactt 1676
             | || |
Db   1053 TCGACCT 1047


RESULT  10
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             2.8%;  Score 50.8;  DB 1;  Length 7218;
  Best Local Similarity   3.3%;  Pred. No. 1.4e-05;
  Matches   13;  Conservative  220;  Mismatches  157;  Indels    0;  Gaps    0;

Qy    211 gaccctgaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatc 270
          || | || | |:  :  :: :: :: :  ::    :: :::: :    :::: :: : ::
Db   1056 GAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1115

Qy    271 cgcatgacccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaag 330
          : : :  ::: : :::::::: ::: :: :  ::::::::  : : : :   :::::
Db   1116 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1175

Qy    331 cccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctctttt 390
          :::::::: ::: :   : :  ::::: :: ::::: : ::  :  :: : ::::::::
Db   1176 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1235

Qy    391 ctccgcagcattacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaat 450
          :::: :  : :: :::::    :::  :    ::   ::: ::  :::   : ::::  :
```

```
Db   1236 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1295

Qy    451 gttggtgaagactgtcccgtctttgacggcctttattccttttgccagacctatgctgga 510
          ::  :    :: :::: :::::  :  ::::: :::::::: ::   ::: : ::
Db   1296 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1355

Qy    511 ggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgg 570
            ::: ::  :  :::: ::   :::  :: :  ::::: :  : :: :: ::  ::
Db   1356 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1415

Qy    571 gctggtggtctccatcacgctaagaagtgc 600
           ::  :  ::::: :: :  ||  || | |
Db   1416 YYYYYYYYYYYYYYYYYYGTACCAAATTC 1445


RESULT  11
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350003
;  CURRENT APPLICATION NUMBER: US/09/007,005B
;  CURRENT FILING DATE: 1998-01-14
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17


  Query Match             2.5%;  Score 44.6;  DB 4;  Length 289;
  Best Local Similarity   7.5%;  Pred. No. 0.00014;
  Matches   20;  Conservative  114;  Mismatches  131;  Indels    0;  Gaps    0;

Qy   1385 attcagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaagagaag 1444
          ::  : :|::: ::::  : :   :     :| | | |  : : : ::  : : : : ::
Db     15 rururarcrurarurururarcrararururarcrararurgrnrnrsrnrnrsrnrnrs 74

Qy   1445 ctgttgaaccagatacaaaggacaaggatggactgaaaggaattatggagcgtggaaaag 1504
             :  ::   ::: : : : :   : :  ::   : : : :   : :::    :: : ::
Db     75 rnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrs 134

Qy   1505 gttgtgaggtggaggtggatgagagtggaagcactaaggttacaggagtaaacccagtgg 1564
           :    :: : ::: : ::: : : : : : : :   : :     : ::  :  :  : ::
Db    135 rnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrs 194

Qy   1565 gagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaacaagggtggggcggagc 1624
           : : ::: : ::  :  : : :   :|:|: : :|: : : : :|: ::: :   : :|
Db    195 rnrnrsrnrnrsrnrnrsrnrnrsrcrargrcrurgrcrgrurararcrurcrururgrc 254

Qy   1625 aggcgtttcctcctaaaacataaga 1649
           :|:|::        |||| | || |
Db    255 rgrcruaaaaaaaaaaaaaaaaaa 279


RESULT  12
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350007
;  CURRENT APPLICATION NUMBER: US/09/244,796
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  EARLIER APPLICATION NUMBER: 09/007,005
;  EARLIER FILING DATE: 1998-01-14
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17


  Query Match             2.5%;  Score 44.6;  DB 4;  Length 289;
  Best Local Similarity   7.5%;  Pred. No. 0.00014;
  Matches   20;  Conservative  114;  Mismatches  131;  Indels    0;  Gaps    0;

Qy   1385 attcagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaagagaag 1444
          ::  : :|::: ::::  : :   :     :| | | |  : : : ::  : : : : ::
Db     15 rururarcrurarurururarcrararururarcrararurgrnrnrsrnrnrsrnrnrs 74

Qy   1445 ctgttgaaccagatacaaaggacaaggatggactgaaaggaattatggagcgtggaaaag 1504
             :  ::   ::: : : : :   : :  ::   : : : :   : :::    :: : ::
Db     75 rnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrs 134

Qy   1505 gttgtgaggtggaggtggatgagagtggaagcactaaggttacaggagtaaacccagtgg 1564
           :    :: : ::: : ::: : : : : : : :   : :     : ::  :  :  : ::
Db    135 rnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrs 194

Qy   1565 gagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaacaagggtggggcggagc 1624
           : : ::: : ::  :  : : :   :|:|: : :|: : : : :|: ::: :   : :|
Db    195 rnrnrsrnrnrsrnrnrsrnrnrsrcrargrcrurgrcrgrurararcrurcrururgrc 254

Qy   1625 aggcgtttcctcctaaaacataaga 1649
           :|:|::        |||| | || |
Db    255 rgrcruaaaaaaaaaaaaaaaaaa 279


RESULT  13
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350003
;  CURRENT APPLICATION NUMBER: US/09/007,005B
;  CURRENT FILING DATE: 1998-01-14
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
```

```
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17


  Query Match             2.1%;  Score 38.8;  DB 4;  Length 289;
  Best Local Similarity   8.6%;  Pred. No. 0.011;
  Matches   19;  Conservative   90;  Mismatches  111;  Indels    0;  Gaps    0;

Qy     197 tttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatga 256
           ||| || ||| |    :|:      :  |:|  : :|  : :| |:| |  : : ::
Db     273 TTTTTTTTTTTTTAYGYCYGYCYAYAYGYAYGYTYTYAYCYGYCYAYGYCYTYGYSYNYN 214

Qy     257 agccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcata 316
           :: :   : : ::: : ::  :     : ::: : : : :  :  : ::: :  :: :
Db     213 YSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYN 154

Qy     317 tgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacg 376
           :::   : : :  :: : ::: : ::: :   : :   : :  :: : ::: : ::  :
Db     153 YSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYN 94

Qy     377 actatgtctcttttctccgcagcattacccctgaaaccca 416
           :: : ::: : : : : :  : : :   : :     :|:|
Db      93 YSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYCYA 54


RESULT  14
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350007
;  CURRENT APPLICATION NUMBER: US/09/244,796
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  EARLIER APPLICATION NUMBER: 09/007,005
;  EARLIER FILING DATE: 1998-01-14
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17


  Query Match             2.1%;  Score 38.8;  DB 4;  Length 289;
  Best Local Similarity   8.6%;  Pred. No. 0.011;
  Matches   19;  Conservative   90;  Mismatches  111;  Indels    0;  Gaps    0;

Qy     197 tttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatga 256
           ||| || ||| |    :|:      :  |:|  : :|  : :| |:| |  : : ::
Db     273 TTTTTTTTTTTTTAYGYCYGYCYAYAYGYAYGYTYTYAYCYGYCYAYGYCYTYGYSYNYN 214

Qy     257 agccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcata 316
           :: :   : : ::: : ::  :     : ::: : : : :  :  : ::: :  :: :
Db     213 YSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYN 154

Qy     317 tgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacg 376
           :::   : : :  :: : ::: : ::: :   : :   : :  :: : ::: : ::  :
Db     153 YSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYN 94

Qy     377 actatgtctcttttctccgcagcattacccctgaaaccca 416
           :: : ::: : : : : :  : : :   : :     :|:|
Db      93 YSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYCYA 54


RESULT  15
US-09-177-325-4
; Sequence 4, Application US/09177325B
; Patent No. 6214983
; GENERAL INFORMATION:
;  APPLICANT:      Tureci, Ozlem
;  APPLICANT:  Sakin, Ugur
;  APPLICANT:  Pfreundschuh, Michael
;  TITLE OF INVENTION:      Isolated Nucleic Acid Molecules Which Encode SCP Proteins
;  TITLE OF INVENTION:  Thereof
;  FILE REFERENCE:      LUD 5525
;  CURRENT APPLICATION NUMBER: US/09/177,325B
;  CURRENT FILING DATE:  1998-10-22
;  NUMBER OF SEQ ID NOS:      15
; SEQ ID NO 4
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
US-09-177-325-4


  Query Match             2.1%;  Score 38.8;  DB 4;  Length 485;
  Best Local Similarity   60.4%;  Pred. No. 0.016;
  Matches   64;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy    1286 agcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcccgaggttg 1345
           || || |   ||||        | ||||||||  ||||| |||||||||    ||||  |
Db     134 aggatccagtaagtcatgagcatgaggaaagaaaacctgttacagagacaaaggagggag 193

Qy    1346 atgaagaccaagaagatggggataaaagatgggatccggattcaga 1391
           ||| |    | ||  |||||||| | | |||    | | ||  ||||
Db     194 atgtaactgatgagcatggggaaagaggatctttgctgaaacaga 239



Search completed: April 28, 2002, 19:24:41
Job time: 8221 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 28, 2002, 18:29:16 ; Search time 203.64 Seconds
(without alignments)
15235.044 Million cell updates/sec

Title: US-09-645-337-1
Perfect score: 1807
Sequence: 1 agagagcagctcccttcccc.....atgatgatatgatatgacaa 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1807	100.0	1807	22	AAF80350	Nucleotide sequenc
	2	830.2	45.9	1805	21	AAZ58260	Soybean histone de
	3	740.4	41.0	1943	20	AAX90840	Maize histone deac
	4	734.2	40.6	1990	21	AAZ58259	Rice histone deace
	5	698.6	38.7	2019	20	AAX90839	Maize histone deac
	6	559.4	31.0	1826	20	AAX90837	Maize histone deac
	7	525	29.1	1609	21	AAC46043	Arabidopsis thalia
	8	525	29.1	1611	21	AAC33022	Arabidopsis thalia
	9	523.4	29.0	1800	22	AAF80351	Nucleotide sequenc
	10	508.8	28.2	1997	24	ABI99512	Mouse ischaemic co
	11	504.4	27.9	1985	20	AAZ32066	Human METH2 relate
	12	504.4	27.9	1985	22	AAC90323	U31814 cDNA clone.
	13	504.4	27.9	1985	22	AAC89555	Human histone deac
	14	499.8	27.7	1449	18	AAT86371	cDNA encoding a hi
	15	499.8	27.7	1611	22	AAC89554	Human histone deac
	16	499.8	27.7	2163	21	AAF16188	Human prostate can
	17	498.2	27.6	2111	17	AAT12940	RPDL transcription
	18	495.8	27.4	2145	23	ABL01897	Drosophila melanog
	19	468	25.9	1576	20	AAX90841	Maize histone deac
	20	451.8	25.0	1954	22	AAC89556	Human histone deac
c	21	421.4	23.3	4755	23	ABL01896	Drosophila melanog
	22	414.8	23.0	1582	23	AAS81302	DNA encoding novel
	23	384.2	21.3	1475	20	AAX90838	Maize histone deac
	24	381.8	21.1	1662	23	ABL04121	Drosophila melanog
c	25	330.2	18.3	3793	23	ABL04120	Drosophila melanog
	26	311.8	17.3	2009	22	AAD17396	Human bone marrow
c	27	255.6	14.1	3826	23	ABL01928	Drosophila melanog
	28	212.6	11.8	1682	22	AAH25122	Nucleotide sequenc
	29	211	11.7	1367	22	AAH75178	Nucleotide sequenc
	30	211	11.7	1654	21	AAZ93329	Human histone deac
	31	211	11.7	2046	22	AAI59530	Human polynucleoti
	32	186	10.3	541	21	AAZ58261	Wheat histone deac
	33	170.4	9.4	402	22	AAH55632	Human breast tumou
	34	166.6	9.2	379	18	AAT86372	Partial cDNA encod
	35	150.2	8.3	375	18	AAT86373	Partial cDNA encod
	36	143	7.9	1539	22	AAI61316	Human polynucleoti
	37	134.6	7.4	1016	22	AAS41223	cDNA encoding nove
	38	131.6	7.3	570	21	AAF07540	Fusarium venenatum
	39	121.8	6.7	241	22	AAS47345	Human breast cance
	40	121.8	6.7	241	22	AAF17915	Human breast cance
	41	105.6	5.8	227	18	AAT86374	Partial cDNA encod
	42	99.6	5.5	222	22	AAC89707	Human ovarian canc
	43	97.6	5.4	411	22	AAS39024	Novel human diagno
c	44	96.4	5.3	286	23	AAS57741	cDNA #417 encoding
	45	89.4	4.9	399	22	AAF67319	Novel human polynu

ALIGNMENTS

RESULT 1
AAF80350
ID AAF80350 standard; DNA; 1807 BP.
XX
AC AAF80350;
XX
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of a histone deacetylase designated AtRPD3A.
XX
KW Histone deacetylase; AtRPD3A; RPD3; gene expression; transgenic plant;
KW HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 142..1647
FT /*tag= a
FT /product= "histone deacetylase AtRPD3A"
XX
PN CA2316036-A1.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX

DR WPI; 2001-258457/27.
DR P-PSDB; AAB67811.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT -
XX
PS Claim 9; Fig 1A; 91pp; English.
XX
CC The present sequence encodes an Arabidopsis thaliana histone deacetylase
CC designated AtRPD3A. The protein is homologous to yeast RPD3 and HDA1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operative
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.
XX
SQ Sequence 1807 BP; 472 A; 386 C; 464 G; 485 T; 0 other;

Query Match 100.0%; Score 1807; DB 22; Length 1807;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1807; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 agagagcagctcccttcccctcggcgaggaggaaggaagaagaaagccagagagagagag 60
Db       1 agagagcagctcccttcccctcggcgaggaggaaggaagaagaaagccagagagagagag 60

Qy      61 agagatcatcgcagcttctcctccgaccatttgactgcgactgtgattacaacacaccgt 120
Db      61 agagatcatcgcagcttctcctccgaccatttgactgcgactgtgattacaacacaccgt 120

Qy     121 tgatcctacgaaaaagaggtaatggatactggcggcaattcgctggcgtccggacctgat 180
Db     121 tgatcctacgaaaaagaggtaatggatactggcggcaattcgctggcgtccggacctgat 180

Qy     181 ggtgtgaagaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggc 240
Db     181 ggtgtgaagaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggc 240

Qy     241 caaggtcatcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactac 300
Db     241 caaggtcatcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactac 300

Qy     301 ggtctccttcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgc 360
Db     301 ggtctccttcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgc 360

Qy     361 cgcttccacgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaa 420
Db     361 cgcttccacgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaa 420

Qy     421 gatcagattcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggc 480
Db     421 gatcagattcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggc 480

Qy     481 ctttattccttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaac 540
Db     481 ctttattccttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaac 540

Qy     541 cacggcctctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgc 600
Db     541 cacggcctctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgc 600

Qy     601 gaggcctctggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaag 660
Db     601 gaggcctctggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaag 660

Qy     661 cagcatgagcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggag 720
Db     661 cagcatgagcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggag 720

Qy     721 gcattttatgctactgacagggttatgactgtctcgtttcataaatttggtgattacttt 780
Db     721 gcattttatgctactgacagggttatgactgtctcgtttcataaatttggtgattacttt 780

Qy     781 cccggtacaggtcacattcaggatataggttatggtagcggaaagtactattctctcaat 840
Db     781 cccggtacaggtcacattcaggatataggttatggtagcggaaagtactattctctcaat 840

Qy     841 gtaccactggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatg 900
Db     841 gtaccactggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatg 900

Qy     901 gggaaagttatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactcc 960
Db     901 gggaaagttatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactcc 960

Qy     961 ctatctggggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtc 1020
Db     961 ctatctggggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtc 1020

Qy    1021 aaatttatgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatc 1080
Db    1021 aaatttatgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatc 1080

Qy    1081 cgcaatgttgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaa 1140
Db    1081 cgcaatgttgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaa 1140

Qy    1141 gacaagatgccggagcatgaatattatgaatactttggtccagactatacacttcacgtt 1200
Db    1141 gacaagatgccggagcatgaatattatgaatactttggtccagactatacacttcacgtt 1200

Qy    1201 gctccaagtaacatggaaaataagaattctcgtcagatgcttgaagagattcgcaatgac 1260
Db    1201 gctccaagtaacatggaaaataagaattctcgtcagatgcttgaagagattcgcaatgac 1260

Qy    1261 cttctccacaatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagacca 1320
Db    1261 cttctccacaatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagacca 1320

Qy    1321 cctgatacagagactcccgaggttgatgaagaccaagaagatggggataaaagatgggat 1380
Db    1321 cctgatacagagactcccgaggttgatgaagaccaagaagatggggataaaagatgggat 1380

Qy    1381 ccggattcagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaaga 1440
Db    1381 ccggattcagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaaga 1440

Qy    1441 gaagctgttgaaccagatacaaaggacaaggatggactgaaaggaattatggagcgtgga 1500
Db    1441 gaagctgttgaaccagatacaaaggacaaggatggactgaaaggaattatggagcgtgga 1500

Qy    1501 aaaggttgtgaggtggaggtggatgagagtggaagcactaaggttacaggagtaaaccca 1560
Db    1501 aaaggttgtgaggtggaggtggatgagagtggaagcactaaggttacaggagtaaaccca 1560

Qy    1561 gtgggagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaacaagggtggggcg 1620
Db    1561 gtgggagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaacaagggtggggcg 1620

Qy    1621 gagcaggcgtttcctcctaaaacataagactcggagcttctaatttcttgctacttttttc 1680
Db    1621 gagcaggcgtttcctcctaaaacataagactcggagcttctaatttcttgctacttttttc 1680

Qy    1681 tgtctatcaaatgttgctagttaagtttctggagttgttgttgttgtaagcactcctctg 1740
```

```
Db   1681 tgtctatcaaatgttgctagttaagtttctggagttgttgttgttgtaagcactcctctg 1740

Qy   1741 ttttagaggattgagcacggatatgtatttattcgttgcatgtctgaatgatgatatgat 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ttttagaggattgagcacggatatgtatttattcgttgcatgtctgaatgatgatatgat 1800

Qy   1801 atgacaa 1807
          |||||||
Db   1801 atgacaa 1807


RESULT   2
AAZ58260
ID   AAZ58260 standard; cDNA; 1805 BP.
XX
AC   AAZ58260;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Soybean histone deacetylase gene 1 (HD1) cDNA clone sr1.pk0023.d1.
XX
KW   Chromatin associated protein; histone deacetylase gene 1; HD1;
KW   soybean; transgenic plant; transcription regulation; ss.
XX
OS   Glycine max.
XX
FH   Key             Location/Qualifiers
FT   CDS             130..1551
FT                   /*tag=  a
XX
PN   WO200004177-A1.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15807.
XX
PR   14-JUL-1998;   98US-0092841.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Vollmer SJ;
XX
DR   WPI; 2000-182439/16.
DR   P-PSDB; AAY58829.
XX
PT   New nucleic acid fragment useful as probes and primers, for
PT   transforming plants  -
XX
PS   Claim 3; Page 26-27; 36pp; English.
XX
CC   The present sequence is that of cDNA clone sr1.pk0023.d1 encoding
CC   soybean histone deacetylase 1 (HD1, see AAY58829), a chromatin
CC   associated protein.  The cDNA clone was isolated from a soybean
CC   root cDNA library on the basis of homology to other plant histone
CC   deacetylases.  The invention relates to isolated rice, soybean and
CC   wheat nucleic acid fragments encoding HD1.  It also relates to the
CC   construction of a chimeric gene encoding all or a portion of HD1,
CC   in sense or antisense orientation, where expression of the chimeric
CC   gene results in production of altered levels of HD1 in a
CC   transformed host cell.  The availability of nucleic acid sequences
CC   encoding (portions) of histone deacetylase proteins will facilitate
CC   studies of global transcriptional regulation in eukaryotic cells,
CC   and will also provide mechanisms to control transcriptional gene
CC   regulation in plants.
XX
SQ   Sequence 1805 BP; 524 A; 309 C; 436 G; 534 T; 2 other;


  Query Match             45.9%;  Score 830.2;  DB 21;  Length 1805;
  Best Local Similarity   76.1%;  Pred. No. 4.4e-244;
  Matches 1021;  Conservative    1;  Mismatches  319;  Indels    0;  Gaps    0;

Qy    128 acgaaaaagaggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtga 187
          |||||||    |||||||| | ||| || || || ||  | || ||  | ||||||||||
Db    116 acgaaaactgagtaatggaaagtggagggaactcccttccatcagggtcagatggtgtga 175

Qy    188 agaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtc 247
          |||| || |||| |||||||||| || ||||| || || || || ||||| || || |
Db    176 agagaaaggtttcatatttctatgatccagaggttggaaactattattatgggcagggac 235

Qy    248 atcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcc 307
          | || ||||| || ||  | || || ||||| ||||| ||  | || ||||| ||  | |
Db    236 acccaatgaaaccacacaggattcgaatgacacatgctcttttagcccactatggattgc 295

Qy    308 ttcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttcc 367
          |||| || |||||||||||| ||||| | ||||    || | || || |||   ||||
Db    296 ttcaacacatgcaggttctgaagcctatggctgctaaagatagggacctttgcaagttcc 355

Qy    368 acgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcaga 427
          | || || || |||||  | |||||  |  |||| ||||||||||| ||||||||||
Db    356 atgctgatgattatgtggcctttctgagaggcatcacccctgaaacgcagcaagatcaat 415

Qy    428 ttcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttatt 487
          |  | || || ||| | || |||||||| |||||||| || || ||||| || ||||| |
Db    416 tgagacagctgaagaggtttaatgttggcgaagactgccctgtatttgatggtctttact 475

Qy    488 ccttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcc 547
          | || |||||||| ||||| ||||| |||||||||||  || | ||| | ||||| ||
Db    476 ctttctgccagacatatgcaggaggttctgttggtggtgctctaaagttgaaccatggag 535

Qy    548 tctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcct 607
          | || |||||||| || || |||||||||||||| ||||| || |||||||| ||||| |
Db    536 tatgtgatattgcaataaattgggctggtggtctacatcatgcaaagaagtgtgaggctt 595

Qy    608 ctggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatg 667
          |||| || || || || ||||| || ||  | |||||  | || || || ||    ||||
Db    596 ctgggttttgctatgttaatgacattgtgctggctattttggaacttctcaaaatacatg 655

Qy    668 agcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttt 727
          |||||||||| ||||| || |||||||||||||| || ||||| || |||||||| ||||
Db    656 agcgtgttctgtatgtggacattgatatccaccatggtgatggtgtagaggaggccttttt 715

Qy    728 atgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggta 787
          |  | || || ||||| |||||||| |||||||||| ||||| |||||||||| || |
Db    716 acaccaccgatagggtcatgactgtttcgtttcataagtttggggattactttcctggaa 775

Qy    788 caggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccac 847
          ||||| | ||||  ||||| || |||| ||  || || || ||||| || ||||| ||
Db    776 caggtgatattcgtgatattggatatgctaaagggaaaatattattcactaaatgttccct 835

Qy    848 tggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaag 907
          |||||||||||||| ||||||||||||||||||    || || ||||| || ||||| || |
Db    836 tggatgatggaattgatgatgagagctatcagtccttgtttaagccaataatgggaaagg 895

Qy    908 ttatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctg 967
          ||||||| |||||  | || || ||||| ||||| ||||||||||||||||||  ||||||
Db    896 ttatggagatttttaggcccggtgctgttgtattacaatgtggtgctgactctttatctg 955

Qy    968 gggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaattta 1027
          ||||  |||||||||| |||||||||| || |||||||||| ||||| |||| || ||
Db    956 gggacaggttaggttgtttcaatctttccataaaaggtcatgcagagtgtgtcagatata 1015

Qy    1028 tgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaatg 1087
           ||||||| || |||||||||| ||  |  | || || ||||| || || ||  | ||||
Db    1016 tgagatcttttaatgttcccctttctattgctcgggggaggtggctatacaataagaaatg 1075

Qy    1088 ttgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaagacaaga 1147
           | || ||||| ||||| | ||||||| | ||||| |||||  ||||| | || || ||||
Db    1076 tggcacgttgttggtgtttcgagactagcgttgctcttgggattgaactagatgataaga 1135

Qy    1148 tgccggagcatgaatattatgaatactttggtccagactatacacttcacgttgctccaa 1207
           ||||  | |||||||||||||||||| |||||||| |||||||| ||||| ||||||||||
Db    1136 tgcctcaacatgaatattatgaatattttggtcctgactatactcttcatgttgctccaa 1195

Qy    1208 gtaacatggaaaatgaagaattctcgtcagatgcttgaagagattcgcaatgaccttctcc 1267
           |||||||||||||| ||||| || || ||  |  | || || ||  |     | ||||||
```

```
Db   1196 gtaacatggaaaacaagaactcccgacaattattggatgaaataagagcaaaacttcttg 1255

Qy   1268 acaatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgata 1327
          | ||| | |||| |||||| ||||| |||||||| ||||| |||||| | |||||||||
Db   1256 ataatttatctaggcttcaacatgcaccaagtgtcccattccaggaacggccacctgatg 1315

Qy   1328 cagagactcccgaggttgatgaagaccaagaagatggggataaaagatgggatccggatt 1387
          ||||||  ||  |||   |||||||| ||||| ||| | ||| ||||||||||||| ||||
Db   1316 cagagcttctagagagagatgaagatcaagatgatagagatgaaagatgggatcccgatt 1375

Qy   1388 cagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaagagaagctg 1447
          | |||| :|| |||| |||||| | || ||| | |  || || || ||||| |||   |
Db   1376 ctgacagrgaggttggcgatgacagcaatcctgttcgcagaagggtgaaaagtgaatgcg 1435

Qy   1448 ttgaaccagatacaaaggaca 1468
          ||||  | ||    || || |
Db   1436 ttgatgctgaggataaagata 1456


RESULT   3
AAX90840
ID   AAX90840 standard; DNA; 1943 BP.
XX
AC   AAX90840;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-4 DNA.
XX
KW   Maize histone deacetylase; family 1, ZmHD1; promoter regulator; promoter;
KW   RNA polymerase II; transcription; plant transformation; heterochromatin;
KW   disease resistance; chromatin assembly; gene activity; toxin screening;
KW   pathogenicity; disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             51..1610
FT                   /*tag= a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28800.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 60-63; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase. This DNA
CC   belongs to family 1, ZmHD1 and appears to be regulator of promoters for
CC   RNA polymerase II, for transcription of genes. The nucleotide sequence
CC   can be used to transform plants and increase disease resistance by
CC   altering the levels of heterochromatin, altering chromatin assembly, and
CC   gene activity of the transformed plants. Additionally, compositions find
CC   use in screening for toxins that affect pathogenicity and in determining
CC   which disease response promoters are regulated by histone deacetylase.
XX
SQ   Sequence 1943 BP; 510 A; 498 C; 497 G; 438 T; 0 other;

  Query Match             41.0%;  Score 740.4;  DB 20;  Length 1943;
  Best Local Similarity   75.0%;  Pred. No. 1.9e-216;
  Matches  953;  Conservative    0;  Mismatches  311;  Indels    6;  Gaps    2;

Qy    144 ggatactggcggcaattcgctggcgtc---cggacctgatggtgtgaagaggaaagtttg 200
          ||   | ||||||||| || ||  ||||   ||| || || ||   |||| ||   || ||
Db     74 gggctccggcggcaactccctcccgtccgtcggccccgacgggcagaagcggcgcgtgtg 133

Qy    201 ttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaagcc 260
          || ||||| ||||| || || ||||| |||||||| || || || ||||| ||||||||
Db    134 ctacttctacgacccggatgtgggcaactactactacgggcagggccatccgatgaagcc 193

Qy    261 ccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatatgca 320
          || |||||||| ||||| ||  | || || || | |||||| |||||  | || |||||
Db    194 gcaccgcatccggatgacgcactcgctgctggcgcgctacggcctcctcaaccagatgca 253

Qy    321 ggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgacta 380
          |||   |   |||  ||| |||||||| || || ||||||||||||||||||||||||||
Db    254 ggtgtaccgccccaacccggcccgcgaccgcgacctctgccgcttccacgccgacgacta 313

Qy    381 tgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaacttaa 440
            ||   || || |||  | | || || ||||| ||||| || ||||| ||||  || ||
Db    314 catcaacttcctgcgctccgtcacgccggaaacgcagcaggaccagatccgcctgctcaa 373

Qy    441 gcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagac 500
          ||||||||| || || || ||||| |||||||| |||||||| ||   ||| ||||||||
Db    374 gcgcttcaacgtcggcgaggactgccccgtcttcgacggcctctacagcttctgccagac 433

Qy    501 ctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgc 560
          |||||| || |  || || || ||  | |||||||| ||||||||||     || || ||
Db    434 ctatgcgggcgcctccgtcggcggggccgtcaagctcaaccacggcc---atgacatcgc 490

Qy    561 catcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgtta 620
          |||||||| | || || || || ||||| ||||||||||||| || ||||||| ||
Db    491 aatcaactggtcgggggggcctgcaccacgccaagaagtgcgaggcgtcgggcttctgcta 550

Qy    621 cgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttcttta 680
          ||||||||| |||||  | || || || ||||| || ||||| || ||| | ||||| ||
Db    551 cgtcaatgacatcgtgctcgccatactcgagctgctcaagcatcacgagagagttctgta 610

Qy    681 tgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactgacag 740
          |||||||||| |||||||||||| || ||||||||||||||||| || ||  | || || ||
Db    611 tgtcgatatcgatatccaccatggtgatggagtggaggaggctttctacacaacagatag 670

Qy    741 ggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattca 800
          |||||||||||||||||| || || |||||||||| || || || |||||  | || |
Db    671 ggttatgactgtctcgttccacaagtttggtgattatttcccaggaacaggggatatccg 730

Qy    801 ggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaat 860
          || || ||  |    |  || |||||||| || || ||||| || || |||||||| ||
Db    731 tgacattgggcactcaaaagggaagtactactccctgaatgtccctctagatgatgggat 790

Qy    861 cgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaatttt 920
          ||||||||| ||||| ||     | || ||||| |||||||| ||||||||||  ||||
Db    791 tgatgatgaaagctaccagtccctttttaagccaatcatgggcaaagttatggaggtttt 850

Qy    921 ccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttagg 980
          ||| || || || || ||  | || |||||||||| ||| | |||||||||  |||| ||
Db    851 ccgccctggtgcagttgtgcttcagtgtggtgctgattccttgtctggggataggttggg 910

Qy    981 ttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaa 1040
          |||||||| || |||||||||||||||| || || || |  | |||||| || |||||
Db    911 ctgcttcaacctctcaatcaaaggtcatgcggaatgtgttaggtatatgaggtctttcaa 970

Qy   1041 tgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctg 1100
          |||||  |  ||||| | |||||||||||| || || ||  | |||||||| || || ||
Db    971 cgttccattgttgctccttggtggtggtggatataccataagaaatgttgcacgctgttg 1030

Qy   1101 gtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatga 1160
          ||| || ||||||||||||||| |||||    ||   |||||||||||||| |  ||||
Db   1031 gtgttatgagactggagttgctcttggccaagagcctgaagacaagatgcctgttaatga 1090

Qy   1161 atattatgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaa 1220
```

```
Db   1091 gtactatgaatacttcggtccagattacactcttcatgttgcaccaagtaacatggagaa 1150

Qy   1221 taagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaa 1280
Db   1151 caaaaatacacgacaacaactggatgatatacgatctaaacttctggataatctttcaaa 1210

Qy   1281 gcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcccga 1340
Db   1211 acttcgacatgctcctagtgtccactttcaagagagagttcctgacacagaaatacctga 1270

Qy   1341 ggttgatgaagaccaagaagatggggataaaagatgggatccggattcagacatggatgt 1400
Db   1271 gcaagatgaagatcaagatgatccagatgaacgacatgatcctgactctgatatggaagt 1330

Qy   1401 tgatgatgac 1410
Db   1331 ggatgaccac 1340
```

RESULT 4
AAZ58259
ID AAZ58259 standard; cDNA; 1990 BP.
XX
AC AAZ58259;
XX
DT 08-MAY-2000 (first entry)
XX
DE Rice histone deacetylase gene 1 (HD1) cDNA contig.
XX
KW Chromatin associated protein; histone deacetylase gene 1; HD1;
KW rice; transgenic plant; transcription regulation; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 3..1484
FT /*tag= a
FT /partial
XX
PN WO200004177-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15807.
XX
PR 14-JUL-1998; 98US-0092841.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Vollmer SJ;
XX
DR WPI; 2000-182439/16.
DR P-PSDB; AAY58828.
XX
PT New nucleic acid fragment useful as probes and primers, for
PT transforming plants -
XX
PS Claim 3; Page 24; 36pp; English.
XX
CC The present sequence is that of a contig encoding a portion
CC (see AAY58828) of rice histone deacetylase 1 (HD1), a chromatin
CC associated protein. The contig was obtained from cDNA clones
CC isolated from rice callus, 15-day-old leaf and 15-day-old seedling
CC cDNA libraries on the basis of homology to other plant histone
CC deacetylases. The invention relates to isolated rice, soybean and
CC wheat nucleic acid fragments encoding HD1. It also relates to the
CC construction of a chimeric gene encoding all or a portion of HD1,
CC in sense or antisense orientation, where expression of the chimeric
CC gene results in production of altered levels of HD1 in a
CC transformed host cell. The availability of nucleic acid sequences
CC encoding (portions) of histone deacetylase proteins will facilitate
CC studies of global transcriptional regulation in eukaryotic cells,
CC and will also provide mechanisms to control transcriptional gene
CC regulation in plants.
XX
SQ Sequence 1990 BP; 502 A; 494 C; 518 G; 476 T; 0 other;

Query Match 40.6%; Score 734.2; DB 21; Length 1990;
Best Local Similarity 76.0%; Pred. No. 1.5e-214;
Matches 920; Conservative 0; Mismatches 288; Indels 3; Gaps 1;

```
Qy    200 gttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaaagc 259
Db      1 gctacttctacgacgcggaggtggggaactactactacgggcaggggcacccgatgaagc 60

Qy    260 cccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatatgc 319
Db     61 cgcaccgcatccggatgacccacgcgctgctcgcccactacggcctcctcgaccagatgc 120

Qy    320 aggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgact 379
Db    121 aggtgctcaagccccacccggcgcgcgaccgcgacctctgccgcttccacgccgacgact 180

Qy    380 atgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaactta 439
Db    181 acgtcgccttcctccgctccgtcacgccggagacccagcaggaccagatccgggcgctca 240

Qy    440 agcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccaga 499
Db    241 agcgcttcaacgtcggcgaggactgccccgtcttcgacggcctctacagcttctgccaga 300

Qy    500 cctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattg 559
Db    301 cctacgccgggggatccgtcggcggcgccgtcaagctcaaccacggcc---acgacatcg 357

Qy    560 ccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgtt 619
Db    358 ccatcaactgggccggcggcctccaccacgccaagaagtgcgaggcctcgggattctgct 417

Qy    620 acgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttcttt 679
Db    418 acgtcaacgacatcgtcctcgccatcctcgagctcctcaaataccaccagcgtgttctct 477

Qy    680 atgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactgaca 739
Db    478 atgtggatatcgatatccaccatggggatggtgtggaggaggcgttctacacgacggaca 537

Qy    740 gggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattc 799
Db    538 gggtgatgacggtctcgttccacaagtttggggattatttcccggggaccggggacattc 597

Qy    800 aggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaa 859
Db    598 gcgatattgggcactcaaaggggaagtattactctctgaatgtcccgttggacgacggta 657

Qy    860 tcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaattt 919
Db    658 tcgacgacgagagctaccagtcgttgttcaagccgatcatggggaaggtgatggaggttt 717

Qy    920 tccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttag 979
Db    718 ttcgccctggcgcggtggtgctccagtgcggtgcggactctctgtcgggtgataggttgg 777

Qy    980 gttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttca 1039
Db    778 gttgcttcaacctgtcaatcaggggccacgcggaatgcgtgagattcatgaggtccttca 837

Qy   1040 atgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgct 1099
Db    838 atgtcccgctgttgctgcttggtggtggtgggtataccataagaaatgttgcgcggtgtt 897

Qy   1100 ggtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatg 1159
Db    898 ggtgctatgagacaggagttgcacttggtcatgagctcactgacaagatgcctccaaatg 957
```

```
Qy   1160 aatattatgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaa 1219
          | |||| ||| ||||||||||||||| |||||||||| ||||| ||||||||||||||| |
Db    958 agtattttgagtactttggtccagattatacacttcatgttgcaccaagtaacatggaga 1017

Qy   1220 ataagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctcta 1279
          | || ||  | || ||| | | || || ||  |       |||||  | ||||| || |
Db   1018 acaaaaacacacgccagcagttggatgatataagatcaagacttcttgataatctttcaa 1077

Qy   1280 agcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcccg 1339
          | ||||  |||||||| || || | |||||| ||  |||| |||||  | |||   || |
Db   1078 aacttcgacatgctcctagcgtccaatttcaagagcgacccccctgaggctgagctacctg 1137

Qy   1340 aggttgatgaagaccaagaagatggggataaaagatgggatccggattcagacatggatg 1399
          ||   |||||||||||||| ||| |||| |||| |||||     || | ||||| ||  ||||
Db   1138 agcaagatgaagaccaagaggatcctgatgaaaggcaccatgctgattctgatgtggaaa 1197

Qy   1400 ttgatgatgac 1410
          | ||||||| |
Db   1198 tggatgatgtc 1208
```

RESULT 5
AAX90839
ID AAX90839 standard; DNA; 2019 BP.
XX
AC AAX90839;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-3 DNA.
XX
KW Maize histone deacetylase; family 1, ZmHD1; promoter regulator; promoter;
KW RNA polymerase II; transcription; plant transformation; heterochromatin;
KW disease resistance; chromatin assembly; gene activity; toxin screening;
KW pathogenicity; disease response promoter; ds.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 140..1459
FT /*tag= a
FT /product= "Maize histone deacetylase"
XX
PN WO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
DR WPI; 1999-611038/52.
DR P-PSDB; AAY28799.
XX
PT New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance -
XX
PS Claim 1; Page 55-58; 87pp; English.
XX
CC The present sequence encodes a maize histone deacetylase. This DNA
CC belongs to family 1, ZmHD1 and appears to be regulator of promoters for
CC RNA polymerase II, for transcription of genes. The nucleotide sequence
CC can be used to transform plants and increase disease resistance by
CC altering the levels of heterochromatin, altering chromatin assembly, and
CC gene activity of the transformed plants. Additionally, compositions find
CC use in screening for toxins that affect pathogenicity and in determining
CC which disease response promoters are regulated by histone deacetylase.
XX

SQ Sequence 2019 BP; 510 A; 532 C; 528 G; 449 T; 0 other;

Query Match 38.7%; Score 698.6; DB 20; Length 2019;
Best Local Similarity 73.0%; Pred. No. 1.3e-203;
Matches 912; Conservative 0; Mismatches 334; Indels 3; Gaps 1;

```
Qy    162 gctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgaggt 221
          || | |  | ||  | || ||   |||| |    || || || ||||| ||| | |||||
Db    175 gccgaccactggcgcggacgggtcgaagcgccgcgtctgctacttctacgacgcggaggt 234

Qy    222 cggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgaccca 281
          ||||| |||||||| || || || || || |||||||| || |||||||||||||||||
Db    235 gggcaactactactacgggcagggccacccgatgaagccgcaccgcatccgcatgaccca 294

Qy    282 tgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagcccttccctgc 341
          || || ||||  | |||||| |||||  | || ||||| ||  ||   ||   ||||||
Db    295 cgcgctgctcggccgctacggcctcctcgaccagatgcaagtgttccgccctcaccctgc 354

Qy    342 ccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagcat 401
          |||||| || || ||||||||||||||||||||||||| || ||||| || |||||   | |
Db    355 ccgcgaccgcgacctctgccgcttccacgccgacgattacgtctccttcctccggtccgt 414

Qy    402 tacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaaaga 461
          ||||| ||||| ||||| || ||||| |||   || ||||||||||| || || || ||
Db    415 cacccccgaaacgcagcaggaccagatccgcgcgctcaagcgcttcaacgtcggcgagga 474

Qy    462 ctgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgttgg 521
          ||| |||||||| ||||| || ||    || || ||||| || || || || ||||||||
Db    475 ctgccccgtcttcgacggtctctacagtttctgtcagacgtacgcgggggctctgttgg 534

Qy    522 tggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggtct 581
          ||| | |||||||| ||||| ||||     ||||| ||||||||||||||| || || ||
Db    535 cggcgccgtcaagctcaaccatggccat---gatatcgccatcaactgggccggcggact 591

Qy    582 ccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtcttagc 641
          ||| ||||| |||||||||||||||||| || || || || || ||||| || ||| | ||
Db    592 ccaccacgccaagaagtgcgaggcctccgggttttgctatgttaatgacattgtcctcgc 651

Qy    642 tatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccacca 701
          ||||| |||||||| ||| | ||  |||| ||||| || || || ||||||||||||||
Db    652 catcctcgagctcctcaagtaccaccagcgcgttctgtacgtggacattgatatccacca 711

Qy    702 cggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgtttca 761
          |||||| || |||||||||||| |||||| | || ||| |||| ||||| ||||| || ||
Db    712 cggggacggcgtggaggaggctttttataccacagaccgggtgatgacagtctcattcca 771

Qy    762 taaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagcgg 821
          || ||||| ||||| || || || ||||| ||||||  ||| | ||  |    |  ||
Db    772 caagtttggagattatttccctgggacaggggacattcgtgatgttgggcactcaaaggg 831

Qy    822 aaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcatct 881
          || || || || || ||||| || ||||| ||||| || |||||||||||||| ||
Db    832 taaatattactccctgaatgttcccctggacgatggtattgatgatgagagctaccagtc 891

Qy    882 gttattcaagcccatcatggggaaagttatggaaatttccgaccaggggctgtggtatt 941
          ||| |||||||| || ||||| || || |||||  | |||   || || || || ||  |
Db    892 gttgttcaagccaataatgggcaaggtgatggaggtcttcaaccctggtgcagtcgtgct 951

Qy    942 gcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaa 1001
          || |||||||| || ||  | || || ||  |||| || || ||||| || || || ||
Db    952 ccagtgtggtgcggattcattgtcgggtgacaggttgggctgtttcaacctctctattaa 1011

Qy   1002 aggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttggg 1061
          ||||| || || || || | |||||||| || ||||| || || ||  ||||  | ||
Db   1012 gggtcacgcagaatgtgtaagatttatgaggtccttcaacgtcccgctgttgctgcttgg 1071

Qy   1062 tggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagttgc 1121
          ||||||||| || || ||  | || ||||| || || |||||||||||||| ||||||||
Db   1072 tggtggtgggtataccataagaaacgttgcacggtgttggtgctacgagacaggagttgc 1131

Qy   1122 acttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggtcc 1181
          |||||   |||  |    |||||||||||     |||| || ||||| || ||||||||
```

```
Db   1132 ccttggtcatgagctcactgacaagatgccacctaatgagtactatgagtattttggtcc 1191

Qy   1182 agactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatgct 1241
Db   1192 agattacactctacatgtcgctccaagtaacatggagaataaaaacacacggcatcaatt 1251

Qy   1242 tgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagtgt 1301
Db   1252 ggatgacataaaatcaaaacttctagataatctttcaaaactccgacatgctcctagtgt 1311

Qy   1302 accatttcaggaaagaccacctgatacagagactcccgaggttgatgaagaccaagaaga 1361
Db   1312 tcagtttcaagagcgacctcctgaggctgagttacctgagcaagatgaagacaaagagaa 1371

Qy   1362 tggggataaaagatgggatccggattcagacatggatgttgatgatgac 1410
Db   1372 tcctgatgaaagacatgatgctgattctgatgtggagatgaatgatgcc 1420


RESULT   6
AAX90837
ID   AAX90837 standard; DNA; 1826 BP.
XX
AC   AAX90837;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-1 DNA.
XX
KW   Maize histone deacetylase; family 1, ZmHD1; promoter regulator; promoter;
KW   RNA polymerase II; transcription; plant transformation; heterochromatin;
KW   disease resistance; chromatin assembly; gene activity; toxin screening;
KW   pathogenicity; disease response promoter; ds.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             29..1405
FT                   /*tag= a
FT                   /product= "Maize histone deacetylase"
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   P-PSDB; AAY28797.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 46-49; 87pp; English.
XX
CC   The present sequence encodes a maize histone deacetylase. This DNA
CC   belongs to family 1, ZmHD1 and appears to be regulator of promoters for
CC   RNA polymerase II, for transcription of genes. The nucleotide sequence
CC   can be used to transform plants and increase disease resistance by
CC   altering the levels of heterochromatin, altering chromatin assembly, and
CC   gene activity of the transformed plants. Additionally, compositions find
CC   use in screening for toxins that affect pathogenicity and in determining
CC   which disease response promoters are regulated by histone deacetylase.
XX
SQ   Sequence 1826 BP; 437 A; 471 C; 475 G; 443 T; 0 other;

  Query Match            31.0%;  Score 559.4;  DB 20;  Length 1826;
  Best Local Similarity  66.8%;  Pred. No. 7.7e-161;
  Matches  797;  Conservative    0;  Mismatches  396;  Indels    0;  Gaps    0;

Qy    160 tcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgag 219
Db     62 tctccggcgggcggggaggatgcgcaccgccgccgcgtcagctatttctacgagccgtcg 121

Qy    220 gtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgacc 279
Db    122 atcggagactactactacgggcaaggtcacccgatgaagccccaccgcatccgaatggcg 181

Qy    280 catgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagccctccct 339
Db    182 cactcgctggtggtccactacggcctccaccgcctcctcgagctctcccgccccctacccg 241

Qy    340 gcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagc 399
Db    242 gcctctgaggccgacatccgccgcttccactccgacgactacgtcgctttcctcgcgtcc 301

Qy    400 attacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaa 459
Db    302 gccaccggaaacccgggtgtcctcgacccgcgcgccattaagcgctttaacgtcggtgag 361

Qy    460 gactgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgtt 519
Db    362 gactgccccgtgttcgacggtctcttcccctctgccaggcctccgctgggggaagcatc 421

Qy    520 ggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggt 579
Db    422 ggcgccgccgtcaagcttaaccgcggggacgccgacatcaccgtcaactgggcgggcggc 481

Qy    580 ctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtctta 639
Db    482 ctccaccacgccaagaagagcgaggcctccgggttctgctacgtcaacgacatcgtcctc 541

Qy    640 gctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccac 699
Db    542 gccatcctcgagctcctcaagttccacaggcgtgtgctatatgtagacattgatgtccac 601

Qy    700 cacggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgttt 759
Db    602 catggagatggcgtggaggaggccttcttcactacaaaccgagtcatgactgtttccttt 661

Qy    760 cataaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagc 819
Db    662 cacaagtatgggggatttttccctggtactggacatatcactgacgttggggcagccgaa 721

Qy    820 ggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcat 879
Db    722 gggaagcattatgctctgaatgttcccctgagtgatggtatcgatgacaccacctttcgt 781

Qy    880 ctgttattcaagcccatcatggggaaagttatggaaattttccgaccaggggctgtggta 939
Db    782 ggtctgtttcaatgcatcattaagaaagttatggaggtttatcagccagacgtggttgtc 841

Qy    940 ttgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatc 999
Db    842 ctccaatgcggagctgactctttggctggagacaggttaggttgcttcaacctgtctgtg 901

Qy   1000 aaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttg 1059
Db    902 aagggtcatgctgactgcctccgtttccttaggtcgtacaatgttcctatgatggtttta 961

Qy   1060 ggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagtt 1119
Db    962 ggtggtggaggttacaccatcagaaatgttgcacgctgctggtgctacgagaccgcagtt 1021

Qy   1120 gcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggt 1179
Db   1022 gctgttggagttgaacctgataacaagctgccttacaatgattactatgagtactttggc 1081

Qy   1180 ccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatg 1239
```

```
          || || ||||| ||||| |    |||| |   | ||||   |||| |     ||
Db   1082 cctgattatactcttcatatccaaccaaaaagtgttgaaaacctgaataccacaaaggac 1141

Qy   1240 cttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagt 1299
          | ||  | ||    ||     |  |  | || || || ||| |  | |||| ||| ||
Db   1142 ttggagaacataaagaacatgatattggagaacctgtcaaagatagaacatgttcccagc 1201

Qy   1300 gtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaaga 1352
             | ||| || || |||||  | ||  | ||  |||| ||||   | || ||
Db   1202 actcaattccatgacagaccgtcagaccctgaagctccagaggagaaagagga 1254


RESULT   7
AAC46043
ID   AAC46043 standard; DNA; 1609 BP.
XX
AC   AAC46043;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 48704.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
```

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.1%; Score 525; DB 21; Length 1609;
Best Local Similarity 66.0%; Pred. No. 2.6e-150;
Matches 778; Conservative 0; Mismatches 395; Indels 6; Gaps 1;

```
Qy   152 gcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatg 211
Db   114 gcggcatctctctgccgtcgggacccgacggacgtaagcggcgagtcagttacttctacg 173

Qy   212 accctgaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatcc 271
Db   174 agccgacgatcggagactactactacggtcaaggccacccgatgaagcctcaccggatcc 233

Qy   272 gcatgacccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagc 331
Db   234 gtatggctcatagcctaatcattcactatcacctccaccgtcgcttagaaatcagtcgcc 293

Qy   332 ccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttc 391
Db   294 ctagcctcgctgacgcctccgatatcggccgattccattcgccggagtatgttgacttcc 353

Qy   392 tccgcagcattacccctgaaacccagcaagatc------agattcgccaacttaagcgct 445
Db   354 tcgcttccgtttcgccggaatctatgggcgatccttccgctgcacgaaacctaaggcgat 413

Qy   446 tcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagacctatg 505
Db   414 tcaatgtcggtgaggattgtcctgtcttcgacggacttttgatttttgccgtgcttccg 473

Qy   506 ctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgccatca 565
Db   474 ccggaggttctattggtgctgccgtcaaattaaacagacaggacgctgatatcgctatca 533

Qy   566 actgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgttacgtca 625
Db   534 attggggcggtgggcttcaccatgctaagaaaagcgaggcttctgggttttgctatgtaa 593

Qy   626 atgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttctttatgtcg 685
Db   594 acgacatcgtgctagggattctggagttgctcaagatgtttaagcgggttctctacatag 653

Qy   686 atattgatatccaccacggggatggagtggaggaggcattttatgctactgacagggtta 745
Db   654 atattgatgtccaccatggagatggagtggaagaagcgttttacaccactgatagagtta 713

Qy   746 tgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattcaggata 805
Db   714 tgactgtttctttccacaaatttggggacttttcccaggaactggtcacataagagatg 773

Qy   806 taggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaatcgatg 865
Db   774 ttggcgctgaaaaagggaaatactatgctctaaatgttccactaaacgatggtatggacg 833

Qy   866 atgagagctatcatctgttattcaagcccatcatggggaaagttatggaaatttccgac 925
Db   834 atgaaagtttccgcagcttgtttagacctcttatccagaaggttatggaagtgtatcagc 893

Qy   926 caggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttaggttgct 985
Db   894 cagaggcagttgttcttcagtgtggtgctgactccttaagtggtgatcggttgggttgct 953

Qy   986 tcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttc 1045
Db   954 tcaacttatcagtcaagggtcacgctgattgccttcggttcttaagatcttacaacgttc 1013

Qy  1046 ccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctggtgct 1105
Db  1014 ctctcatggtgttgggtggtggagggtatactattcgaaatgttgcccgttgctggtgtt 1073

Qy  1106 acgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatgaatatt 1165
Db  1074 atgagactgcagttgctgttggagtagagccggacaacaaactcccttacaatgagtatt 1133

Qy  1166 atgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaataaga 1225
Db  1134 ttgagtatttcggcccagattatacgcttcatgtcgacccaagtcctatggagaatttaa 1193

Qy  1226 attctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaagcttc 1285
```

```
          |  | |   |     | ||   |||  | ||     | ||  |  | || ||    ||
Db   1194 acacgcccaaagatatggagaggataaggaacacgttgctggaacaactttcgggactaa 1253

Qy   1286 agcatgctccaagtgtaccatttcaggaaagaccacctg 1324
            || || || || || |  |||||| | | |||||| |
Db   1254 tacacgcacctagcgtccagtttcagcacacaccaccag 1292


RESULT   8
AAC33022
ID   AAC33022 standard; DNA; 1611 BP.
XX
AC   AAC33022;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 1498.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
```

```
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             29.1%;  Score 525;  DB 21;  Length 1611;
  Best Local Similarity   66.0%;  Pred. No. 2.6e-150;
  Matches  778;  Conservative    0;  Mismatches  395;  Indels    6;  Gaps    1;

Qy     152 gcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatg 211
           ||||||  || ||| |||| ||||| || ||    ||| ||  |||  |||| ||||| |
Db     116 gcggcatctctctgccgtcgggacccgacggacgtaagcggcgagtcagttacttctacg 175

Qy     212 accctgaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatcc 271
           | ||   | ||||  | |||||||| || ||||| || || |||||||| || || ||||
Db     176 agccgacgatcggagactactactacggtcaaggccacccgatgaagcctcaccggatcc 235

Qy     272 gcatgacccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagc 331
           | ||| | |||  |||  ||  ||||||    ||||  |  |   |  |  |      |
Db     236 gtatggctcatagcctaatcattcactatcacctccaccgtcgcttagaaatcagtcgcc 295

Qy     332 ccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctctttc 391
           |   ||  ||   ||     ||| || |||| ||||  |    || |||||    || |
Db     296 ctagcctcgctgacgcctccgatatcggccgattccattcgccggagtatgttgacttcc 355

Qy     392 tccgcagcattacccctgaaacccagcaagatc------agattcgccaacttaagcgct 445
           ||     | || | || ||| |   |   ||||          ||  | || | ||| |
Db     356 tcgcttccgtttcgccggaatctatgggcgatccttccgctgcacgaaacctaaggcgat 415

Qy     446 tcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagacctatg 505
           ||||||| ||||| || ||||| ||||| ||||| |||| |   |||||||   | |  |
Db     416 tcaatgtcggtgaggattgtcctgtcttcgacggacttttgatttttgccgtgcttccg 475

Qy     506 ctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgccatca 565
           | ||||| ||| ||||||   | |||||  | |||        |   |||||  || ||||
Db     476 ccggaggttctattggtgctgccgtcaaattaaacagacaggacgctgatatcgctatca 535

Qy     566 actgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgttacgtca 625
           | ||||  ||||| || || || ||||||||  ||||||| ||||| || || || | |
Db     536 attggggcggtgggcttcaccatgctaagaaaagcgaggcttctgggttttgctatgtaa 595

Qy     626 atgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttctttatgtcg 685
           | || |||||  |||  || || ||| | || |||  |  | |||| ||||| ||  | |
Db     596 acgacatcgtgctaggaattctggagttgctcaagatgtttaagcgggttctctacatag 655

Qy     686 atattgatatccaccacggggatggagtggaggaggcattttatgctactgacagggtta 745
           |||||||| ||||||| || |||||||||||| || || |||||  | ||||| || ||||
Db     656 atattgatgtccaccatggagatggagtggaagaagcgttttacaccactgatagagtta 715

Qy     746 tgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattcaggata 805
           ||||||| || || || |||||||| || |  || || || || ||||||||    |||
Db     716 tgactgtttctttccacaaatttggggacttttcccaggaactggtcacataagagatg 775

Qy     806 taggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaatcgatg 865
           | ||   ||  |  || || |||||| |||| ||||| |||||  | ||||| || || |
Db     776 ttggcgctgaaaaagggaaatactatgctctaaatgttccactaaacgatggtatggacg 835

Qy     866 atgagagctatcatctgttattcaagcccatcatggggaaagttatggaaattttccgac 925
           |||| || |  |     || || |  ||  | ||   ||| |||||||||| | |  |  |
Db     836 atgaaagtttccgcagcttgtttagacctcttatccagaaggttatggaagtgtatcagc 895

Qy     926 caggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttaggttgct 985
           ||| ||| || ||  | || ||||||||||||||| ||  ||| |||||||| |||||||
Db     896 cagaggcagttgttcttcagtgtggtgctgactccttaagtggtgatcggttgggttgct 955

Qy     986 tcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttc 1045
           ||||  | ||| |||| ||||| ||||| ||| |    ||  | ||||| | ||| ||||
Db     956 tcaacttatcagtcaagggtcacgctgattgccttcggttcttaagatcttacaacgttc 1015

Qy    1046 ccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctggtgct 1105
           | ||  || | ||||||||||| || || ||||| || ||||||||||||||||||| |
Db    1016 ctctcatggtgttgggtggtggagggtatactattcgaaatgttgcccgttgctggtgtt 1075

Qy    1106 acgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatgaatatt 1165
           | ||||||| |||||  |||||| ||    ||  ||||  | ||  |  |||| ||||
Db    1076 atgagactgcagttgctgttggagtagagccggacaacaaactcccttacaatgagtatt 1135

Qy    1166 atgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaataaga 1225
            ||| || || || ||||| ||||| ||||| || |  ||||||   |||||  |||   |
Db    1136 ttgagtatttcggcccagattatacgcttcatgtcgacccaagtcctatggagaatttaa 1195

Qy    1226 attctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaagcttc 1285
           |  | |   |     | ||   |||  | ||     | ||  |  | || ||    ||
Db    1196 acacgcccaaagatatggagaggataaggaacacgttgctggaacaactttcgggactaa 1255

Qy    1286 agcatgctccaagtgtaccatttcaggaaagaccacctg 1324
```

```
Db   1256 tacacgcacctagcgtccagtttcagcacacaccaccag 1294


RESULT   9
AAF80351
ID   AAF80351 standard; DNA; 1800 BP.
XX
AC   AAF80351;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Nucleotide sequence of a histone deacetylase designated AtRPD3B.
XX
KW   Histone deacetylase; AtRPD3B; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation; ds.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             121..1536
FT                   /*tag= a
FT                   /product= "histone deacetylase AtRPD3B"
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
XX
PF   24-AUG-2000; 2000CA-2316036.
XX
PR   27-AUG-1999;   99US-0383971.
XX
PA   (MIAC ) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI   Miki B,  Brown D,  Tian L,  Wu K;
XX
DR   WPI; 2001-258457/27.
DR   P-PSDB; AAB67812.
XX
PT   Methods for regulating gene expression in transgenic plants, e.g.
PT   repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT   elongation), comprises introducing genes encoding histone deacetylase
PT   -
XX
PS   Claim 9; Fig 1B; 91pp; English.
XX
CC   The present sequence encodes an Arabidopsis thaliana histone deacetylase
CC   designated AtRPD3B. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence 1800 BP; 478 A; 390 C; 456 G; 476 T; 0 other;


  Query Match             29.0%;  Score 523.4;  DB 22;  Length 1800;
  Best Local Similarity   65.9%;  Pred. No. 8.8e-150;
  Matches  777;  Conservative    0;  Mismatches  396;  Indels    6;  Gaps    1;

Qy     152 gcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatg 211
Db     137 gcggcatctctctgccgtcgggacccgacggacgtaagcggcgagtcagttacttctacg 196

Qy     212 accctgaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatcc 271

Db     197 agccgacgatcggagactactactacggtcaaggccacccgatgaagcctcaccggatcc 256

Qy     272 gcatgacccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagc 331
Db     257 gtatggctcatagcctaatcattcactatcacctccaccgtcgcttagaaatcagtcgcc 316

Qy     332 ccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctctttc 391
Db     317 ctagcctcgctgacgcctccgatatcggccgattccattcgccggagtatgttgacttcc 376

Qy     392 tccgcagcattacccctgaaacccagcaagatc------agattcgccaacttaagcgct 445
Db     377 tcgcttccgtttcgccggaatctatgggcgatccttccgctgcacgaaacctaaggcgat 436

Qy     446 tcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagacctatg 505
Db     437 tcaatgtcggtgaggattgtcctgtcttcgacggacttttgattttgccgtgcttccg 496

Qy     506 ctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgccatca 565
Db     497 ccggaggttctattggtgctgccgtcaaattaaacagacaggacgctgatatcgctatca 556

Qy     566 actgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgttacgtca 625
Db     557 attggggcggtgggcttcaccatgctaagaaaagcgaggcttctgggttttgctatgtaa 616

Qy     626 atgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttctttatgtcg 685
Db     617 acgacatcgtgctagggattctggagttgctcaagatgtttaagcgggttctctacatag 676

Qy     686 atattgatatccaccacggggatggagtggaggaggcattttatgctactgacagggtta 745
Db     677 atattgatgtccaccatggagatggagtggaagaagcgttttacaccactgatagagtta 736

Qy     746 tgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattcaggata 805
Db     737 tgactgtttctttccacaaatttggggacttttcccaggaactggtcacataagagatg 796

Qy     806 taggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaatcgatg 865
Db     797 ttggcgctgaaaaagggaaatactatgctctaaatgttccactaaacgatggtatggacg 856

Qy     866 atgagagctatcatctgttattcaagcccatcatggggaaagttatggaaattttccgac 925
Db     857 atgaaagtttccgcagcttgtttagacctcttatccagaaggttatggaagtgtatcagc 916

Qy     926 caggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttaggttgct 985
Db     917 cagaggcagttgttcttcagtgtggtgctgactccttaagtggtgatcggttgggttgct 976

Qy     986 tcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttc 1045
Db     977 tcaacttatcagtcaagggtcacgctgattgccttcggttcttaagatcttacaacgttc 1036

Qy    1046 ccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctggtgct 1105
Db    1037 ctctcatggtgttgggtggtgaagggtatactattcgaaatgttgcccgttgctggtgtt 1096

Qy    1106 acgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatgaatatt 1165
Db    1097 atgagactgcagttgctgttggagtagagccggacaacaaactcccttacaatgagtatt 1156

Qy    1166 atgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaataaga 1225
Db    1157 ttgagtatttcggcccagattatacgcttcatgtcgacccaagtcctatggagaatttaa 1216

Qy    1226 attctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaagcttc 1285
Db    1217 acacgcccaaagatatggagaggataaggaacacgttgctggaacaacttttcgggactaa 1276

Qy    1286 agcatgctccaagtgtaccatttcaggaaagaccacctg 1324
```

```
Db   1277 tacacgcacctagcgtccagtttcagcacacaccaccag 1315


RESULT  10
ABI99512
ID   ABI99512 standard; cDNA; 1997 BP.
XX
AC   ABI99512;
XX
DT   07-MAR-2002  (first entry)
XX
DE   Mouse ischaemic condition related cDNA sequence SEQ ID NO:493.
XX
KW   Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW   vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
OS   Mus musculus.
XX
PN   WO200188188-A2.
XX
PD   22-NOV-2001.
XX
PF   18-MAY-2001; 2001WO-JP04192.
XX
PR   18-MAY-2000; 2000JP-0145977.
XX
PA   (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI   Ishikawa K,  Asai S,  Takahashi Y,  Nagata T,  Ishii Y;
XX
DR   WPI; 2002-034733/04.
DR   P-PSDB; ABB57202.
XX
PT   Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT   expression levels of particular genes defined in the specification or
PT   by determining the expression profile of a gene group comprising these
PT   genes -
XX
PS   Claim 2; Page 1369-1373; 2690pp; English.
XX
CC   The present invention describes a method for examining ischaemic
CC   conditions, comprising measuring the expression levels of particular
CC   genes (I) in a test sample or determining the expression profile of a
CC   gene group in the sample comprising genes selected from (I). The method
CC   is useful for examining the ischaemic condition (e.g. compressive
CC   ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC   expression levels of particular genes (ABI99202 to ABI99912, encoding
CC   the protein sequences in ABB57020 to ABB57374) or by determining the
CC   expression profile of a gene group comprising these genes. The
CC   expression levels or expression profiles produced by these genes are
CC   used as an indicator when screening for ischaemic condition-improving
CC   drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC   represent PCR primers for a mouse ischaemic condition related sequence,
CC   which are used in the exemplification of the present invention.
XX
SQ   Sequence 1997 BP; 595 A; 376 C; 486 G; 540 T; 0 other;


  Query Match             28.2%;  Score 508.8;  DB 24;  Length 1997;
  Best Local Similarity   64.0%;  Pred. No. 2.9e-145;
  Matches  768;  Conservative    0;  Mismatches  432;  Indels    0;  Gaps    0;

Qy     164 tggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgaggtcg 223
           |||||| | | |  | || | |||| |||||| || || | ||||||   |||  | |
Db     170 tggcgtacagtcaaggaggcggcaagaagaaagtgtgctactactatgatggtgatattg 229

Qy     224 gcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgacccatg 283
           ||||||| || |||||||| ||||||||||||||||| ||| | ||||| ||||| |||
Db     230 gcaattattattatggccagggtcatcccatgaagcctcatagaatccggatgactcata 289

Qy     284 ccctcctcgctcactacggtctccttcagcatatgcaggttctcaagcccttccctgccc 343
           | | ||     | || ||| |    |   | ||| |  |    | |||       |||
Db     290 acttgctgctaaattatggtttataccgaaaaatggaaatatataggcctcataaagcca 349

Qy     344 gcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagcatta 403
           |     ||  |       | ||||  ||| || ||| ||   ||||| ||    || |
Db     350 ctgctgaagaaatgactaaataccacagcgatgagtatatcaagtttctacgatcaataa 409

Qy     404 cccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaagact 463
           || || |    |   ||  | | |   ||  |  || | || || || || ||||| |
Db     410 gaccagataatatgtctgagtacagtaagcagatgcagagatttaacgtcggagaagatt 469

Qy     464 gtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgttggtg 523
           |||| || ||||| || || | |   ||||| |||  ||   | || || || |||| ||
Db     470 gtccggtgtttgatggactctttgagttttgtcagctctccacgggtggttcagttgctg 529

Qy     524 gctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggtctcc 583
           |  |||| ||  | ||||     |     ||||| ||  |||| ||||||||| || || |
Db     530 gggctgtgaaattaaaccggcaacaaactgatatggctgtcaattgggctggaggactac 589

Qy     584 atcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtcttagcta 643
           |||| || |||||||  || || || || ||||| || || |||||||| ||  | || |
Db     590 atcatgccaagaagtcagaagcatcagggttctgctatgttaatgatattgtgcttgcca 649

Qy     644 tcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccaccacg 703
           |||| ||  | ||||||| | ||| || | ||  | ||| | || || || |||||||| |
Db     650 tcctcgaattacttaagtatcatcagagagtcttatatattgacatagacatccaccatg 709

Qy     704 gggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgtttcata 763
           | ||||| || ||||| || |||||| | || ||  | || ||||| ||||| || ||||
Db     710 gtgatggtgttgaggaagctttttatacaacagatcgcgtgatgaccgtctcattccata 769

Qy     764 aatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagcggaa 823
           ||| ||| || |||||||| || |||||  || |   |||||| ||| ||| |  ||||
Db     770 aatatggggaatacttttcctggaacaggagacttgagggatattggtgctggaaagggaa 829

Qy     824 agtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcatctgt 883
           | |||||| || ||||| | ||  ||   ||||| || |||||||||    |||   |
Db     830 aatactatgctgtcaattttcccatgagagatggtatagatgatgaatcatatggacaaa 889

Qy     884 tattcaagcccatcatggggaaagttatggaaattttccgaccaggggctgtggtattgc 943
           | || ||||| |||||    ||||| ||||| || | ||  ||  | || |||||  |||
Db     890 tttttaagcctatcatctcaaaagtgatggagatgtaccagcctagcgcggtggtgctgc 949

Qy     944 aatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaaag 1003
           | ||||| || |||||||| ||||||||  || | ||||| ||||||||  |  ||||||
Db     950 agtgtggcgcagactccctgtctggggacaggcttggttgtttcaatctaactgtcaaag 1009

Qy    1004 gtcatgctgaagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttgggtg 1063
           |||||||| | || ||  || |   || | | || ||  | ||  | ||| |  | ||||
Db    1010 gtcatgctaaatgtgtagaagtagcgaaaacttttaacttgccattgctgatgctcggtg 1069

Qy    1064 gtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagttgcac 1123
           | || || ||||| ||||| |||||||||| || |||   || ||||||| |||||| |
Db    1070 gaggaggctacacaatccggaatgttgcccgatgttggacatatgagactgcagttgccc 1129

Qy    1124 ttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggtccag 1183
           |||    |||| ||    |  || ||||  |  |||| || | ||| || ||||| ||||
Db    1130 ttgattgtgaaattcccaatgagttgccatataatgattacttttgagtattttggaccag 1189

Qy    1184 actatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatgcttg 1243
           |||  | ||| ||  ||  |||    ||||||  |||  ||||  |||   |     | |
Db    1190 acttcaaactgcatattagtccttcaaacatgacaaaccagaacactccagaatatatgg 1249

Qy    1244 aagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagtgtac 1303
           || ||||     |     |  |  | |||||   || | |  |||||  ||  |||| |
Db    1250 aaaagataaaacagcgtttatttgaaaatctacgtatgttaccacatgcacctggtgttc 1309

Qy    1304 catttcaggaaagaccacctgatacagagactcccgaggttgatgaagaccaagaagatg 1363
           | | || |  |  |||   ||| | |    |   ||   ||  || ||  | | ||| |
Db    1310 aaatgcaagctattccagaggatgctgttcatgaagacagtggagatgaggatggagaag 1369


RESULT  11
AAZ32066
ID   AAZ32066 standard; DNA; 1985 BP.
```

```
XX
AC   AAZ32066;
XX
DT   10-JAN-2000  (first entry)
XX
DE   Human METH2 related EST U31814.
XX
KW   Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW   cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW   angiogenesis inhibitor; abnormal wound healing; inflammation;
KW   rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW   diabetic retinopathy; macula degeneration; haemangioma; detection;
KW   arterial-venous malformation; immune deficiency; ss.
XX
OS   Homo sapiens.
XX
PN   WO9937660-A1.
XX
PD   29-JUL-1999.
XX
PF   22-JAN-1999;   99WO-US01313.
XX
PR   23-JAN-1998;   98US-0072298.
PR   28-AUG-1998;   98US-0098539.
XX
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM;
XX
DR   WPI; 1999-590684/50.
XX
PT   New isolated metalloprotease thrombospondin polypeptides, useful for
PT   treating hyperproliferative disorders, cancers or autoimmune disorders
PT   -
XX
PS   Disclosure; Page 444-446; 457pp; English.
XX
CC   AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC   metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC   respectively. METH1 and METH2 have been found to be potent inhibitors of
CC   angiogenesis both in vitro and in vivo. They can be used for treating
CC   cancer and other disorders related to angiogenesis including abnormal
CC   wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC   endometrial bleeding disorders, diabetic retinopathy, some forms of
CC   macula degeneration, haemangiomas, and arterial-venous malformations.
CC   They may be useful in treating deficiencies or disorders of the immune
CC   system, by activating or inhibiting the proliferation, differentiation,
CC   or mobilisation (chemotaxis) of immune cells. The etiology of these
CC   immune deficiencies or disorders may be genetic, somatic, such as
CC   cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC   toxins), or infectious. They can also be used to treat inflammatory
CC   conditions, both chronic and acute conditions. The products can also be
CC   used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC   AAY49511 represent sequences given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;

  Query Match             27.9%;  Score 504.4;  DB 20;  Length 1985;
  Best Local Similarity   63.2%;  Pred. No. 6.5e-144;
  Matches  775;  Conservative    0;  Mismatches  451;  Indels    0;  Gaps    0;

Qy    138 ggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagt 197
          |||   ||     | |||  |  |  |||||| | | |  |  || |  || |  |||||
Db    180 ggtggcggcggtggccggggagcccatggcgtacagtcaaggaggcggcaaaaaaaaagt 239

Qy    198 ttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaa 257
            || || | ||| |||  |||  | || ||||| || ||||| || ||||||||||||||
Db    240 ctgctactactacgacggtgatattggaaattattattatggacagggtcatcccatgaa 299

Qy    258 gccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatat 317
          ||| ||| | |||||||||||||||  | | ||     | || ||  |        | ||
Db    300 gcctcatagaatccgcatgacccataacttgctgttaaattatggcttatacagaaaaat 359

Qy    318 gcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacga 377
          | |  |    | ||||      |||   |      ||  |       |  |||   || ||
Db    360 ggaaatatataggccccataaagccactgccgaagaaatgacaaaatatcacagtgatga 419

Qy    378 ctatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaact 437
          ||| ||   ||||| ||    || |  || || | |  |   ||  | | |   || |
Db    420 gtatatcaaatttctacggtcaataagaccagataacatgtctgagtatagtaagcagat 479

Qy    438 taagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgcca 497
            |   || |||||||| ||||| ||||| |  ||||| || || | |   ||||| ||
Db    480 gcatatatttaatgttggagaagattgtccagcgtttgatggactctttgagttttgtca 539

Qy    498 gacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatat 557
          |  ||   |||| || || |||| |||  |||| ||| | ||||      |     |||||
Db    540 gctctcaactggcggttcagttgctggagctgtgaagttaaaccgacaacagactgatat 599

Qy    558 tgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctg 617
          ||  | || |||||||| ||  | ||||| |||||||| | ||| || || || |||||
Db    600 ggctgttaattgggctggaggattacatcatgctaagaaatacgaagcatcaggattctg 659

Qy    618 ttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttct 677
          ||||||  |||||||| ||  | || ||||| ||  | || ||| | ||| || | ||  |
Db    660 ttacgttaatgatattgtgcttgccatccttgaattactaaagtatcatcagagagtctt 719

Qy    678 ttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactga 737
          ||| | ||||| ||||| || || || ||||| || || || || |||||| | || ||
Db    720 atatattgatatagatattcatcatggtgatggtgttgaagaagctttttatacaacaga 779

Qy    738 cagggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacat 797
          | || ||||| || || || ||||||| ||| || |||||||| || ||||| |  | |
Db    780 tcgtgtaatgacggtatcattccataaatatggggaatactttcctggcacaggagactt 839

Qy    798 tcaggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatgg 857
            |||||| |||  ||| |  || || |||||| || ||||| | ||| ||  ||||||
Db    840 gagggatattggtgctggaaaaggcaaatactatgctgtcaattttccaatgtgtgatgg 899

Qy    858 aatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaat 917
          || |||||||||   |||   | | |||| ||||| || ||    || || ||||| ||
Db    900 tatagatgatgagtcatatgggcagatatttaagcctattatctcaaaggtgatggagat 959

Qy    918 tttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggtt 977
          |  | |||  | |||||||||| || |||||||| |||||  |||||| ||| |  |
Db    960 gtatcaacctagtgctgtggtattacagtgtggtgcagactcattatctggtgatagact 1019

Qy    978 aggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgtt 1037
          ||||| ||||||||  || |||||||||||||| | || ||  || || | | | | ||
Db    1020 gggttgtttcaatctaacagtcaaaggtcatgctaaatgtgtagaagttgtaaaaacttt 1079

Qy    1038 caatgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttg 1097
           ||  | ||  ||||| |  | || || ||||| ||||| ||||| ||||||||  || ||
Db    1080 taacttaccattactgatgcttggaggaggtggctacacaatccgtaatgttgctcgatg 1139

Qy    1098 ctggtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagca 1157
           |||   || ||||||| |||||| ||||     |||  ||    |  || ||||  | |
Db    1140 ttggacatatgagactgcagttgcccttgattgtgagattcccaatgagttgccatataa 1199

Qy    1158 tgaatattatgaatactttggtccagactatacacttcacgttgctccaagtaacatgga 1217
           ||| || | ||| || ||||| |||||||  | ||| ||  ||  |||    ||||||
Db    1200 tgattactttgagtattttggaccagacttcaaactgcatattagtccttcaaacatgac 1259

Qy    1218 aaataagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctc 1277
           ||| ||||  |||   |     | ||| ||||     |     |  |  | ||| |
Db    1260 aaaccagaacactccagaatatatggaaaagataaaacagcgtttgtttgaaaatttgcg 1319

Qy    1278 taagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcc 1337
           | | | |  ||||| ||  |||| |   | || |  |  |||   ||| | |    |
Db    1320 catgttacctcatgcacctggtgtccagatgcaagctattccagaagatgctgttcatga 1379

Qy    1338 cgaggttgatgaagaccaagaagatg 1363
```

```
           ||   ||  || ||  | | ||| |
Db  1380 agacagtggagatgaagatggagaag 1405


RESULT  12
AAC90323
ID   AAC90323 standard; DNA; 1985 BP.
XX
AC   AAC90323;
XX
DT   19-MAR-2001  (first entry)
XX
DE   U31814 cDNA clone.
XX
KW   METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW   cancer therapy; benign tumour; ocular angiogenic disease;
KW   rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW   vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW   scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW   coronary collateral; cerebral collateral; arteriovenous malformation;
KW   ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW   plaque neovascularisation; telangiectasia; haemophiliac joint; EST;
KW   angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW   Crohn's disease; atherosclerosis; birth control; ss.
XX
OS   Unidentified.
XX
PN   WO200071577-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US14462.
XX
PR   25-MAY-1999;   99US-0318208.
PR   20-JUL-1999;   99US-0144882.
PR   10-AUG-1999;   99US-0147823.
PR   13-AUG-1999;   99US-0373658.
PR   22-DEC-1999;   99US-0171503.
PR   22-FEB-2000; 2000US-0183792.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
PA   (JONA/) JONAK Z L.
PA   (TRUL/) TRULLI S H.
PA   (FORN/) FORNWALD J A.
PA   (TERR/) TERRETT J A.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM,  Jonak ZL,  Trulli SH;
PI   Fornwald JA,  Terrett JA;
XX
DR   WPI; 2001-025136/03.
XX
PT   METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT   inhibit angiogensis in the treatment of disorders such as cancer,
PT   rheumatoid arthritis and psoriasis -
XX
PS   Claim 14; Pages 745-746; 768pp; English.
XX
CC   The present invention relates to human METH1 and METH2, (ME for
CC   metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC   The present sequence is an expressed sequence tag (EST) for METH. METH
CC   can be used for inhibiting angiogenesis in an individual, and for
CC   treating cancer, benign tumours, an ocular angiogenic disease,
CC   rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC   vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC   scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC   coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC   ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC   neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC   fibromuscular dysplasia, wound granulation, Crohn's disease or
CC   atherosclerosis. METH can also be used in birth control. METH can also
CC   be used in diagnostic methods for the prognosis of cancer.
XX
SQ   Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;


  Query Match             27.9%;  Score 504.4;  DB 22;  Length 1985;
  Best Local Similarity   63.2%;  Pred. No. 6.5e-144;
  Matches  775;  Conservative    0;  Mismatches  451;  Indels    0;  Gaps    0;

Qy    138 ggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaaaagt 197
          |||   ||     | |||  |  |  |||||| | | |  |  || |  || |  |||||
Db    180 ggtggcggcggtggccggggagcccatggcgtacagtcaaggaggcggcaaaaaaaaagt 239

Qy    198 ttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaa 257
          || || | ||| |||  |||  | || ||||| || ||||| || |||||||||||||||
Db    240 ctgctactactacgacggtgatattggaaattattattatggacagggtcatcccatgaa 299

Qy    258 gccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatat 317
          ||| ||| | ||||||||||||||||  | | ||     | || ||  |        | ||
Db    300 gcctcatagaatccgcatgacccataacttgctgttaaattatggcttatacagaaaaat 359

Qy    318 gcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacga 377
          | |  |    | ||||       |||   |      ||  |        |  |||   || ||
Db    360 ggaaatatataggccccataaagccactgccgaagaaatgacaaaatatcacagtgatga 419

Qy    378 ctatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaact 437
          |||  ||   ||||| ||    || |  || || | |  |   ||  | | |   ||  |
Db    420 gtatatcaaatttctacggtcaataagaccagataacatgtctgagtatagtaagcagat 479

Qy    438 taagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgcca 497
          |    || |||||||| ||||| ||||| |  ||||| || || | | |   ||||| ||
Db    480 gcatatatttaatgttggagaagattgtccagcgtttgatggactctttgagttttgtca 539

Qy    498 gacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatat 557
          |  ||   |||| || || |||| |||  |||| ||| | ||||     |     |||||
Db    540 gctctcaactggcggttcagttgctggagctgtgaagttaaaccgacaacagactgatat 599

Qy    558 tgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctg 617
          ||  | || |||||||| ||  | ||||| |||||||| | ||| || || || |||||
Db    600 ggctgttaattgggctggaggattacatcatgctaagaaatacgaagcatcaggattctg 659

Qy    618 ttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttct 677
          |||||| |||||||| ||  | || |||||  ||  | || ||| | ||| || | ||  |
Db    660 ttacgttaatgatattgtgcttgccatccttgaattactaaagtatcatcagagagtctt 719

Qy    678 ttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactga 737
          ||| | ||||| |||| || || || ||||| || || || || |||||| | || ||
Db    720 atatattgatatagatattcatcatggtgatggtgttgaagaagctttttatacaacaga 779

Qy    738 cagggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacat 797
          | || ||||| || || || ||||||| ||| || |||||||| || |||||  || |
Db    780 tcgtgtaatgacggtatcattccataaatatggggaatactttcctggcacaggagactt 839

Qy    798 tcaggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatgg 857
            |||||| |||  ||| |  || || |||||| || ||||| | ||| ||  ||||||
Db    840 gagggatattggtgctggaaaaggcaaatactatgctgtcaattttccaatgtgtgatgg 899

Qy    858 aatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaat 917
          || ||||||||   |||   | | |||| ||||| || ||    || || ||||| ||
Db    900 tatagatgatgagtcatatgggcagatatttaagcctattatctcaaaggtgatggagat 959

Qy    918 tttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggtt 977
          |  | |||  | |||||||||||| || |||||||| |||||  |||||| ||| |  |
Db    960 gtatcaacctagtgctgtggtattacagtgtggtgcagactcattatctggtgatagact 1019

Qy    978 aggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgtt 1037
          ||||| ||||||||  || ||||||||||||| | || ||  || || | | | ||
Db   1020 gggttgtttcaatctaacagtcaaaggtcatgctaaatgtgtagaagttgtaaaaacttt 1079

Qy   1038 caatgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttg 1097
          ||  | ||  ||||| |  | || || ||||| ||||| ||||| |||||||| || ||
```

```
Db   1080 taacttaccattactgatgcttggaggaggtggctacacaatccgtaatgttgctcgatg 1139

Qy   1098 ctggtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagca 1157
          |||   || |||||||| |||||| ||||    |||  ||    |  || ||||  |  |
Db   1140 ttggacatatgagactgcagttgcccttgattgtgagattcccaatgagttgccatataa 1199

Qy   1158 tgaatattatgaatactttggtccagactatacacttcacgttgctccaagtaacatgga 1217
          ||| || | ||| || ||||| |||||||  | ||| ||  ||  |||    ||||||
Db   1200 tgattactttgagtattttggaccagacttcaaactgcatattagtccttcaaacatgac 1259

Qy   1218 aaataagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctc 1277
          |||  ||||  |||   |     | ||| ||||     |     |  |  | ||| |
Db   1260 aaaccagaacactccagaatatatggaaaagataaaacagcgtttgtttgaaaatttgcg 1319

Qy   1278 taagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcc 1337
          | | | |  ||||| ||  |||| |   | || |  |  |||   ||| | |    |
Db   1320 catgttacctcatgcacctggtgtccagatgcaagctattccagaagatgctgttcatga 1379

Qy   1338 cgaggttgatgaagaccaagaagatg 1363
          ||   ||  || ||  | | ||| |
Db   1380 agacagtggagatgaagatggagaag 1405


RESULT 13
AAC89555
ID   AAC89555 standard; DNA; 1985 BP.
XX
AC   AAC89555;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Human histone deacetylase HDAC-2 coding sequence.
XX
KW   Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
KW   HDAC-D; cell cycle; tumourigenesis; cancer; inhibitor; antisense;
KW   gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200071703-A2.
XX
PD   30-NOV-2000.
XX
PF   03-MAY-2000; 2000WO-IB01252.
XX
PR   03-MAY-1999;   99US-0132287.
XX
PA   (METH-) METHYLGENE INC.
XX
PI   Macleod AR,  Li Z,  Besterman JM;
XX
DR   WPI; 2001-016407/02.
DR   P-PSDB; AAB49955.
XX
PT   Antisense oligonucleotide that inhibits expression of a histone
PT   deacetylase, useful for treating and/or alleviating the symptoms of
PT   neoplasia, or for inhibiting neoplastic cell growth in an animal -
XX
PS   Disclosure; Page 53; 125pp; English.
XX
CC   The present invention provides inhibitors of histone deacetylase enzymes
CC   such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These
CC   inhibitors may be antisense strands or they may be compounds identified
CC   by contacting the enzyme with the compound and measuring the resulting
CC   enzyme activity. These inhibitors are useful for treating cancers and for
CC   identifying which histone deacetylase is involved in a neoplasia.
XX
SQ   Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;

  Query Match           27.9%;  Score 504.4;  DB 22;  Length 1985;
  Best Local Similarity  63.2%;  Pred. No. 6.5e-144;
  Matches  775;  Conservative    0;  Mismatches  451;  Indels    0;  Gaps    0;

Qy    138 ggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagt 197
          |||   ||      | |||  |  |  |||||| | | |  |  || |  || |  |||||
Db    180 ggtggcggcggtggccggggagcccatggcgtacagtcaaggaggcggcaaaaaaaaagt 239

Qy    198 ttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaa 257
          || || | ||| |||  |||  | || |||||  || |||||  || ||||||||||||||
Db    240 ctgctactactacgacggtgatattggaaattattattatggacagggtcatcccatgaa 299

Qy    258 gccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatat 317
          ||| ||| | ||||||||||||||||  | | ||     | || ||  |        | ||
Db    300 gcctcatagaatccgcatgacccataacttgctgttaaattatggcttatacagaaaaat 359

Qy    318 gcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacga 377
          | |  |    | ||||       |||   |      ||  |        |  |||   || ||
Db    360 ggaaatatataggccccataaagccactgccgaagaaatgacaaaatatcacagtgatga 419

Qy    378 ctatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaact 437
          ||| ||   ||||| ||    || |  || || | |  |   ||  | | |   ||  |
Db    420 gtatatcaaatttctacggtcaataagaccagataacatgtctgagtatagtaagcagat 479

Qy    438 taagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgcca 497
          |    || |||||||| ||||| ||||| |  ||||| || || | |   ||||| ||
Db    480 gcatatatttaatgttggagaagattgtccagcgtttgatggactctttgagttttgtca 539

Qy    498 gacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatat 557
          |  ||   |||| || || |||| |||  |||| ||| | ||||     |     |||||
Db    540 gctctcaactggcggttcagttgctggagctgtgaagttaaaccgacaacagactgatat 599

Qy    558 tgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctg 617
          ||  | || |||||||| ||  | |||| |||||||| | ||| || || || |||||
Db    600 ggctgttaattgggctggaggattacatcatgctaagaaatacgaagcatcaggattctg 659

Qy    618 ttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttct 677
          ||||||  |||||||| ||  | || |||||  ||  | || ||| | ||| || | ||  |
Db    660 ttacgttaatgatattgtgcttgccatccttgaattactaaagtatcatcagagagtctt 719

Qy    678 ttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactga 737
          ||| | |||||  |||||  || || || |||||  || || || || |||||| | || ||
Db    720 atatattgatatagatattcatcatggtgatggtgttgaagaagctttttatacaacaga 779

Qy    738 cagggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacat 797
          | || |||||  || || || |||||||  ||| ||  ||||||||  || |||||   || |
Db    780 tcgtgtaatgacggtatcattccataaatatggggaatactttcctggcacaggagactt 839

Qy    798 tcaggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatgg 857
            ||||||  |||  |||  |  || || ||||||  || |||||  | |||  ||  ||||||
Db    840 gagggatattggtgctggaaaaggcaaatactatgctgtcaattttccaatgtgtgatgg 899

Qy    858 aatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaat 917
          || |||||||||   |||   | | ||||  |||||  || ||    || || |||||  ||
Db    900 tatagatgatgagtcatatgggcagatatttaagcctattatctcaaaggtgatggagat 959

Qy    918 tttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggtt 977
          |  | |||  | ||||||||||| || ||||||| |||||  ||||||| ||| |  |
Db    960 gtatcaacctagtgctgtggtattacagtgtggtgcagactcattatctggtgatagact 1019

Qy    978 aggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgtt 1037
          |||||  ||||||||  || ||||||||||||||| | || ||  || || | | | ||
Db   1020 gggttgtttcaatctaacagtcaaaggtcatgctaaatgtgtagaagttgtaaaaacttt 1079

Qy   1038 caatgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttg 1097
          ||  | ||  ||||| |  | || || ||||| ||||| ||||| |||||||| || ||
Db   1080 taacttaccattactgatgcttggaggaggtggctacacaatccgtaatgttgctcgatg 1139

Qy   1098 ctggtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagca 1157
          |||   || |||||| |||| ||||    ||| ||    |  || ||||  |  |
Db   1140 ttggacatatgagactgcagttgcccttgattgtgagattcccaatgagttgccatataa 1199

Qy   1158 tgaatattatgaatactttggtccagactatacacttcacgttgctccaagtaacatgga 1217
          ||| || | ||| || ||||| |||||||  | ||| ||  ||  |||    ||||||
Db   1200 tgattactttgagtattttggaccagacttcaaactgcatattagtccttcaaacatgac 1259
```

```
Qy   1218 aaataagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctc 1277
          |||  ||||  |||   |     | ||| ||||     |     |  |  | |||| |
Db   1260 aaaccagaacactccagaatatatggaaaagataaaacagcgtttgtttgaaaatttgcg 1319

Qy   1278 taagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcc 1337
          | | | |  ||||| ||  |||| |   | || |  |  |||   ||| | |    |
Db   1320 catgttacctcatgcacctggtgtccagatgcaagctattccagaagatgctgttcatga 1379

Qy   1338 cgaggttgatgaagaccaagaagatg 1363
          ||   ||  || ||  | | ||| |
Db   1380 agacagtggagatgaagatggagaag 1405


RESULT  14
AAT86371
ID   AAT86371 standard; cDNA; 1449 BP.
XX
AC   AAT86371;
XX
DT   16-APR-1998  (first entry)
XX
DE   cDNA encoding a histone deacetylase (HDx) protein, designated HD1.
XX
KW   Histone deacetylase gene; HDx; HD1; HDx polypeptide; deacetylation; H3;
KW   H4; cell differentiation; chromatin structure; cell cycle progression;
KW   proliferative disorder; fibroproliferative disorder;
KW   degenerative disorder; autoimmune disease; HDx inhibitor; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..0
FT                   /*tag=  a
XX
PN   WO9735990-A2.
XX
PD   02-OCT-1997.
XX
PF   26-MAR-1997;   97WO-US05275.
XX
PR   26-MAR-1996;   96US-0624735.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Hassig CA,  Jamison TF,  Schreiber SL,  Taunton J;
XX
DR   WPI; 1997-489651/45.
DR   P-PSDB; AAW29324.
XX
PT   New isolated histone deacetylase polypeptide(s) and genes - used to
PT   develop products for modulating the proliferation, survival or
PT   differentiation of cells, e.g. for treating tumours.
XX
PS   Claim 27; Pages 112-114; 159pp; English.
XX
CC   The present cDNA sequence encodes a novel histone deacetylase (HDx)
CC   polypeptide, designated HD1. The HDx polypeptides are capable of
CC   modulating proliferation survival and differentiation of cells. The
CC   proteins are able to alter chromatin structure by deacetylating histones
CC   such as H3 or H4. They have the ability to modulate cell growth by
CC   influencing cell cycle progression or to modulate gene transcription. The
CC   products can be used for diagnosis and therapy. They can be used, for
CC   example, to treat tumours or proliferative disorders or spermatogenesis,
CC   osteogenesis, chondrogenesis or the differentiation of progenitor cells.
CC   They can also be used to treat psoriasis, bone diseases,
CC   fibroproliferative disorders, degenerative disorders, or for repair of
CC   cartilage, increasing bone density, liver repair subsequent to a partial
CC   hepatectomy, to promote regeneration of lung tissue in the treatment of
CC   emphysema, or for inducing tolerance in autoimmune diseases, and
CC   transplant recipients. HDx inhibitors can be used as anti-fungal agents,
CC   preservatives in foodstuff, feed supplements for promoting weight gain in
CC   livestock, disinfectants, insecticides or defoliants.  The products can
CC   also be used in cell cultures.
XX
SQ   Sequence 1449 BP; 406 A; 329 C; 396 G; 318 T; 0 other;


  Query Match             27.7%;  Score 499.8;  DB 18;  Length 1449;
  Best Local Similarity   64.1%;  Pred. No. 1.4e-142;
  Matches  753;  Conservative    0;  Mismatches  422;  Indels    0;  Gaps    0;

Qy    189 gaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtca 248
          ||||||||| ||||| | ||| |||   || || || |||||||| ||||| ||||| ||
Db     24 gaggaaagtctgttactactacgacggggatgttggaaattactattatggacaaggcca 83

Qy    249 tcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcct 308
          || |||||||| || || ||||||||||| |||    | ||     |||| ||||||
Db     84 cccaatgaagcctcaccgaatccgcatgactcataatttgctgctcaactatggtctcta 143

Qy    309 tcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttcca 368
          |   | ||| |  |        ||   |   |||   |     ||  |   |   | |||
Db    144 ccgaaaaatggaaatctatcgccctcacaaagccaatgctgaggagatgaccaagtacca 203

Qy    369 cgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagat 428
          |  ||| |||||  |    ||  | |||  |||    || || | |  |   ||  | |
Db    204 cagcgatgactacattaaattcttgcgctccatccgtccagataacatgtcggagtacag 263

Qy    429 tcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattc 488
          ||  |  || | ||||| |||||||| |||||||| || || || ||||| | |
Db    264 caagcagatgcagagattcaacgttggtgaggactgtccagtattcgatggcctgtttga 323

Qy    489 cttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcct 548
          || || |||   | | |||| || ||||| |   |  |||| || |||||  |    |
Db    324 gttctgtcagttgtctactggtggttctgtggcaagtgctgtgaaacttaataagcagca 383

Qy    549 ctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctc 608
          || || ||  | || |||||||| || || || || || ||||||| |||||| ||
Db    384 gacggacatcgctgtgaattgggctgggggctgcaccatgcaaagaagtccgaggcatc 443

Qy    609 tggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatga 668
          |||||||||||||||||||||||||||||| || ||||| || || || ||| | ||  |
Db    444 tggcttctgttacgtcaatgatatcgtcttggccatcctggaactgctaaagtatcacca 503

Qy    669 gcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttta 728
          | | || || ||  | || |||||||| ||||| || || || ||||| ||||| || ||
Db    504 gagggtgctgtacattgacattgatattcaccatggtgacggcgtggaagaggccttcta 563

Qy    729 tgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtac 788
            | || ||| |||| |||||||| || |||||||| | ||| || ||||| || || ||
Db    564 caccacggaccgggtcatgactgtgtcctttcataagtatggagagtacttcccaggaac 623

Qy    789 aggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccact 848
          ||  || | | |||||| ||   ||| |  || ||||| ||| || | | ||    || ||
Db    624 tggggacctacgggatatcggggctggcaaaggcaagtattatgctgttaactacccgct 683

Qy    849 ggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagt 908
          || || || ||||| ||| |||| |     | ||||||||  |||||   |||||
Db    684 ccgagacgggattgatgacgagtcctatgaggccattttcaagccggtcatgtccaaagt 743

Qy    909 tatggaaatttttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctgg 968
          ||||| || ||||  ||  | || ||||| || || |||||  | ||||||||||||||
Db    744 aatggagatgttccagcctagtgcggtggtcttacagtgtggctcagactccctatctgg 803

Qy    969 ggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttat 1028
          |||||||||||||||||||||||  | |||||||| || ||  |||| ||  ||||| |
Db    804 ggatcggttaggttgcttcaatctaactatcaaaggacacgccaagtgtgtggaatttgt 863

Qy   1029 gagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaatgt 1088
          |     || ||  | ||  | ||| |  |||| || ||||||||||| || || || ||
Db    864 caagagctttaacctgcctatgctgatgctgggaggcggtggttacaccattcgtaacgt 923

Qy   1089 tgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaagacaagat 1148
          |||||| ||||||   || ||||| |  || || || |     ||  |     |  || |
Db    924 tgcccggtgctggacatatgagacagctgtggccctggatacggagatccctaatgagct 983

Qy   1149 gccggagcatgaatattatgaatactttggtccagactatacacttcacgttgctccaag 1208
```

```
Db    984 tccatacaatgactactttgaatactttggaccagatttcaagctccacatcagtccttc 1043

Qy   1209 taacatggaaaatagaattctcgtcagatgcttgaagagattcgcaatgaccttctcca 1268
Db   1044 caatatgactaaccagaacacgaatgagtacctggagaagatcaaacagcgactgtttga 1103

Qy   1269 caatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatac 1328
Db   1104 gaaccttagaatgctgccgcacgcacctggggtccaaatgcaggcgattcctgaggacgc 1163

Qy   1329 agagactcccgaggttgatgaagaccaagaagatg 1363
Db   1164 catccctgaggagagtggcgatgaggacgaagacg 1198


RESULT 15
AAC89554
ID   AAC89554 standard; DNA; 1611 BP.
XX
AC   AAC89554;
XX
DT   08-MAR-2001  (first entry)
XX
DE   Human histone deacetylase HDAC-1 coding sequence.
XX
KW   Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
KW   HDAC-D; cell cycle; tumourigenesis; cancer; inhibitor; antisense;
KW   gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200071703-A2.
XX
PD   30-NOV-2000.
XX
PF   03-MAY-2000; 2000WO-IB01252.
XX
PR   03-MAY-1999;   99US-0132287.
XX
PA   (METH-) METHYLGENE INC.
XX
PI   Macleod AR,  Li Z,  Besterman JM;
XX
DR   WPI; 2001-016407/02.
DR   P-PSDB; AAB49954.
XX
PT   Antisense oligonucleotide that inhibits expression of a histone
PT   deacetylase, useful for treating and/or alleviating the symptoms of
PT   neoplasia, or for inhibiting neoplastic cell growth in an animal -
XX
PS   Disclosure; Page 50-51; 125pp; English.
XX
CC   The present invention provides inhibitors of histone deacetylase enzymes
CC   such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These
CC   inhibitors may be antisense strands or they may be compounds identified
CC   by contacting the enzyme with the compound and measuring the resulting
CC   enzyme activity. These inhibitors are useful for treating cancers and for
CC   identifying which histone deacetylase is involved in a neoplasia.
XX
SQ   Sequence 1611 BP; 428 A; 385 C; 440 G; 358 T; 0 other;


  Query Match           27.7%;  Score 499.8;  DB 22;  Length 1611;
  Best Local Similarity  64.1%;  Pred. No. 1.5e-142;
  Matches  753;  Conservative    0;  Mismatches  422;  Indels    0;  Gaps    0;

Qy    189 gaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtca 248
Db    134 gaggaaagtctgttactactacgacggggatgttggaaattactattatggacaaggcca 193

Qy    249 tcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctcct 308
Db    194 cccaatgaagcctcaccgaatccgcatgactcataatttgctgctcaactatggtctcta 253

Qy    309 tcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttcca 368
Db    254 ccgaaaaatggaaatctatcgccctcacaaagccaatgctgaggagatgaccaagtacca 313

Qy    369 cgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagat 428
Db    314 cagcgatgactacattaaattcttgcgctccatccgtccagataacatgtcggagtacag 373

Qy    429 tcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattc 488
Db    374 caagcagatgcagagattcaacgttggtgaggactgtccagtattcgatggcctgtttga 433

Qy    489 cttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcct 548
Db    434 gttctgtcagttgtctactggtggttctgtggcaagtgctgtgaaacttaataagcagca 493

Qy    549 ctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctc 608
Db    494 gacggacatcgctgtgaattgggctggggcctgcaccatgcaaagaagtccgaggcatc 553

Qy    609 tggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatga 668
Db    554 tggcttctgttacgtcaatgatatcgtcttggccatcctggaactgctaaagtatcacca 613

Qy    669 gcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggaggcatttta 728
Db    614 gagggtgctgtacattgacattgatattcaccatggtgacggcgtggaagaggccttcta 673

Qy    729 tgctactgacagggttatgactgtctcgtttcataaatttggtgattactttcccggtac 788
Db    674 caccacggaccgggtcatgactgtgtcctttcataagtatggagagtacttcccaggaac 733

Qy    789 aggtcacattcaggatataggttatggtagcggaaagtactattctctcaatgtaccact 848
Db    734 tggggacctacgggatatcggggctggcaaaggcaagtattatgctgttaactacccgct 793

Qy    849 ggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagt 908
Db    794 ccgagacgggattgatgacgagtcctatgaggccattttcaagccggtcatgtccaaagt 853

Qy    909 tatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctgg 968
Db    854 aatggagatgttccagcctagtgcggtggtcttacagtgtggctcagactccctatctgg 913

Qy    969 ggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttat 1028
Db    914 ggatcggttaggttgcttcaatctaactatcaaaggacacgccaagtgtgtggaatttgt 973

Qy   1029 gagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaatgt 1088
Db    974 caagagctttaacctgcctatgctgatgctgggaggcggtggttacaccattcgtaacgt 1033

Qy   1089 tgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaagacaagat 1148
Db   1034 tgcccggtgctggacatatgagacagctgtggccctggatacggagatccctaatgagct 1093

Qy   1149 gccggagcatgaatattatgaatactttggtccagactatacacttcacgttgctccaag 1208
Db   1094 tccatacaatgactactttgaatactttggaccagatttcaagctccacatcagtccttc 1153

Qy   1209 taacatggaaaatagaattctcgtcagatgcttgaagagattcgcaatgaccttctcca 1268
Db   1154 caatatgactaaccagaacacgaatgagtacctggagaagatcaaacagcgactgtttga 1213

Qy   1269 caatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatac 1328
Db   1214 gaaccttagaatgctgccgcacgcacctggggtccaaatgcaggcgattcctgaggacgc 1273

Qy   1329 agagactcccgaggttgatgaagaccaagaagatg 1363
Db   1274 catccctgaggagagtggcgatgaggacgaagacg 1308
```

Search completed: April 28, 2002, 19:28:40
Job time: 3564 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 28, 2002, 16:52:30 ; Search time 1866.22 Seconds
(without alignments)
20262.483 Million cell updates/sec

Title: US-09-645-337-1
Perfect score: 1807
Sequence: 1 agagagcagctcccttcccc.....atgatgatatgatatgacaa 1807

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1807	100.0	1807	6	AX139014	AX139014 Sequence
	2	1807	100.0	1807	8	AF195547	AF195547 Arabidops
	3	1807	100.0	1808	8	AF014824	AF014824 Arabidops
c	4	959	53.1	147215	8	ATF20D10	AL035538 Arabidops
c	5	959	53.1	199789	8	ATCHRIV89	AL161593 Arabidops
	6	795	44.0	1839	8	AF282858	AF282858 Mesembrya
	7	740.4	41.0	1906	8	AF384032	AF384032 Zea mays
	8	740.4	41.0	1943	6	AR168366	AR168366 Sequence
	9	714.6	39.5	1579	8	AF332875	AF332875 Oryza sat
	10	709.6	39.3	2047	8	AF035815	AF035815 Zea mays
	11	698.6	38.7	2019	6	AR168365	AR168365 Sequence
	12	559.4	31.0	1742	8	AF045473	AF045473 Zea mays
	13	559.4	31.0	1826	6	AR168363	AR168363 Sequence
	14	557	30.8	1662	8	AF440226	AF440226 Zea mays
	15	525	29.1	1704	8	AY072201	AY072201 Arabidops
	16	523.4	29.0	1733	8	AF195548	AF195548 Arabidops
	17	523.4	29.0	1800	6	AX139016	AX139016 Sequence
	18	508.8	28.2	1997	6	AX305742	AX305742 Sequence
	19	508.8	28.2	1997	10	MMU31758	U31758 Mus musculu
	20	504.4	27.9	1985	6	AX053103	AX053103 Sequence
	21	504.4	27.9	1985	9	HSU31814	U31814 Human trans
	22	502.4	27.8	154409	2	AP003724	AP003724 Oryza sat
	23	500.2	27.7	1977	10	MMHIDE	X98207 M.musculus
	24	499.8	27.7	1611	6	AX053101	AX053101 Sequence
	25	499.8	27.7	1611	9	HSU50079	U50079 Human histo
	26	499.8	27.7	2093	9	BC000301	BC000301 Homo sapi
	27	499.4	27.6	1999	10	MMU80780	U80780 Mus musculu
	28	498.2	27.6	2091	9	D50405	D50405 Human mRNA
	29	498.2	27.6	2111	6	AR012007	AR012007 Sequence
	30	498.2	27.6	2111	6	I62388	I62388 Sequence 2
	31	498.2	27.6	2111	23	E11455	E11455 cDNA encodi
	32	495.8	27.4	2157	3	AY058487	AY058487 Drosophil
	33	492.2	27.2	1646	5	AF043328	AF043328 Gallus ga
	34	492.2	27.2	1646	5	AF044169	AF044169 Gallus ga
	35	492.2	27.2	1678	5	AF039751	AF039751 Gallus ga
	36	492.2	27.2	2305	5	XLAB21	X78454 X.laevis AB
	37	489.4	27.1	1487	5	AF020658	AF020658 Xenopus l
	38	483.2	26.7	2180	3	AF032919	AF032919 Strongylo
	39	479.8	26.6	2172	3	DMHISSEAC	Y09258 D.melanogas
	40	479.6	26.5	2073	3	AF026949	AF026949 Drosophil
	41	474.4	26.3	1878	5	AF039752	AF039752 Gallus ga
	42	473.6	26.2	1694	8	AF440228	AF440228 Zea mays
	43	468	25.9	1576	6	AR168367	AR168367 Sequence
	44	452.4	25.0	1799	10	AF321131	AF321131 Rattus no
	45	451.8	25.0	1954	6	AX053105	AX053105 Sequence

ALIGNMENTS

RESULT 1
AX139014
LOCUS AX139014 1807 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1094112.
ACCESSION AX139014
VERSION AX139014.1 GI:14274698
KEYWORDS .
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 1807)
AUTHORS Wu,K., Miki,B.L., Tian,L. and Brown,D.C.
TITLE Repressing gene expression in plants
JOURNAL Patent: EP 1094112-A 1 25-APR-2001;
The Minister of Agriculture and Agri-Food (CA)
FEATURES Location/Qualifiers
source 1. .1807
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
BASE COUNT 472 a 386 c 464 g 485 t
ORIGIN

```
  Query Match           100.0%;  Score 1807;  DB 6;  Length 1807;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1807;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 agagagcagctccctcccctcggcgaggaggaaggaagaagaaagccagagagagagag 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 AGAGAGCAGCTCCCTTCCCCTCGGCGAGGAGGAAGGAAGAAGAAAGCCAGAGAGAGAG 60

Qy      61 agagatcatcgcagcttctcctccgaccatttgactgcgactgtgattacaacacaccgt 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 AGAGATCATCGCAGCTTCTCCTCCGACCATTTGACTGCGACTGTGATTACAACACACCGT 120

Qy     121 tgatcctacgaaaaagaggtaatggatactggcggcaattcgctggcgtccggacctgat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGATCCTACGAAAAAGAGGTAATGGATACTGGCGGCAATTCGCTGGCGTCCGGACCTGAT 180

Qy     181 ggtgtgaagaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGTGTGAAGAGGAAAGTTTGTTATTTCTATGACCCTGAGGTCGGCAATTACTACTATGGC 240

Qy     241 caaggtcatcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CAAGGTCATCCCATGAAGCCCCATCGCATCCGCATGACCCATGCCCTCCTCGCTCACTAC 300

Qy     301 ggtctccttcagcatatgcaggttctcaagccctccctgcccgcgaacgtgatctctgc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGTCTCCTTCAGCATATGCAGGTTCTCAAGCCCTTCCCTGCCCGCGAACGTGATCTCTGC 360

Qy     361 cgcttccacgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaa 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CGCTTCCACGCCGACGACTATGTCTCTTTTCTCCGCAGCATTACCCCTGAAACCCAGCAA 420

Qy     421 gatcagattcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GATCAGATTCGCCAACTTAAGCGCTTCAATGTTGGTGAAGACTGTCCCGTCTTTGACGGC 480

Qy     481 ctttattccttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaac 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTTATTCCTTTTGCCAGACCTATGCTGGAGGATCTGTTGGTGGCTCTGTCAAGCTTAAC 540

Qy     541 cacggcctctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgc 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CACGGCCTCTGCGATATTGCCATCAACTGGGCTGGTGGTCTCCATCACGCTAAGAAGTGC 600

Qy     601 gaggcctctggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GAGGCCTCTGGCTTCTGTTACGTCAATGATATCGTCTTAGCTATCCTAGAGCTCCTTAAG 660

Qy     661 cagcatgagcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggag 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CAGCATGAGCGTGTTCTTTATGTCGATATTGATATCCACCACGGGGATGGAGTGGAGGAG 720

Qy     721 gcattttatgctactgacagggttatgactgtctcgtttcataaatttggtgattacttt 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GCATTTTATGCTACTGACAGGGTTATGACTGTCTCGTTTCATAAATTTGGTGATTACTTT 780

Qy     781 cccggtacaggtcacattcaggatataggttatggtagcggaaagtactattctctcaat 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CCCGGTACAGGTCACATTCAGGATATAGGTTATGGTAGCGGAAAGTACTATTCTCTCAAT 840

Qy     841 gtaccactggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatg 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GTACCACTGGATGATGGAATCGATGATGAGAGCTATCATCTGTTATTCAAGCCCATCATG 900

Qy     901 gggaaagttatggaaatttttccgaccaggggctgtggtattgcaatgtggtgctgactcc 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GGGAAAGTTATGGAAATTTTCCGACCAGGGGCTGTGGTATTGCAATGTGGTGCTGACTCC 960

Qy     961 ctatctggggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtc 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 CTATCTGGGGATCGGTTAGGTTGCTTCAATCTTTCAATCAAAGGTCATGCTGAGTGCGTC 1020

Qy    1021 aaatttatgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatc 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 AAATTTATGAGATCGTTCAATGTTCCCCTACTGCTCTTGGGTGGTGGTGGTTACACTATC 1080

Qy    1081 cgcaatgttgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaa 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CGCAATGTTGCCCGTTGCTGGTGCTACGAGACTGGAGTTGCACTTGGAGTTGAAGTTGAA 1140

Qy    1141 gacaagatgccggagcatgaatattatgaatactttggtccagactatacacttcacgtt 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GACAAGATGCCGGAGCATGAATATTATGAATACTTTGGTCCAGACTATACACTTCACGTT 1200

Qy    1201 gctccaagtaacatggaaaataagaattctcgtcagatgcttgaagagattcgcaatgac 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 GCTCCAAGTAACATGGAAAATAAGAATTCTCGTCAGATGCTTGAAGAGATTCGCAATGAC 1260

Qy    1261 cttctccacaatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagacca 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CTTCTCCACAATCTCTCTAAGCTTCAGCATGCTCCAAGTGTACCATTTCAGGAAAGACCA 1320

Qy    1321 cctgatacagagactcccgaggttgatgaagaccaagaagatggggataaaagatgggat 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCTGATACAGAGACTCCCGAGGTTGATGAAGACCAAGAAGATGGGGATAAAAGATGGGAT 1380

Qy    1381 ccggattcagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaaga 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 CCGGATTCAGACATGGATGTTGATGATGACCGTAAACCTATACCAAGCAGAGTAAAAAGA 1440

Qy    1441 gaagctgttgaaccagatacaaaggacaaggatggactgaaaggaattatggagcgtgga 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 GAAGCTGTTGAACCAGATACAAAGGACAAGGATGGACTGAAAGGAATTATGGAGCGTGGA 1500

Qy    1501 aaaggttgtgaggtggaggtggatgagagtggaagcactaaggttacaggagtaaacccca 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 AAAGGTTGTGAGGTGGAGGTGGATGAGAGTGGAAGCACTAAGGTTACAGGAGTAAACCCA 1560

Qy    1561 gtgggagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaacaagggtggggcg 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GTGGGAGTGGAGGAAGCAAGTGTGAAAATGGAAGAGGAAGGAACAAACAAGGGTGGGGCG 1620

Qy    1621 gagcaggcgtttcctcctaaaacataagactcggagcttctaatttcttgctacttttttc 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GAGCAGGCGTTTCCTCCTAAAACATAAGACTCGGAGCTTCTAATTTCTTGCTACTTTTTC 1680

Qy    1681 tgtctatcaaatgttgctagttaagtttctggagttgttgttgttgtaagcactcctctg 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 TGTCTATCAAATGTTGCTAGTTAAGTTTCTGGAGTTGTTGTTGTTGTAAGCACTCCTCTG 1740

Qy    1741 ttttagaggattgagcacggatatgtatttattcgttgcatgtctgaatgatgatatgat 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 TTTTAGAGGATTGAGCACGGATATGTATTTATTCGTTGCATGTCTGAATGATGATATGAT 1800

Qy    1801 atgacaa 1807
           |||||||
Db    1801 ATGACAA 1807


RESULT   2
AF195547
LOCUS       AF195547               1807 bp    mRNA    linear   PLN 01-NOV-2000
DEFINITION  Arabidopsis thaliana putative histone deacetylase (RPD3A) mRNA,
            complete cds.
ACCESSION   AF195547
VERSION     AF195547.1  GI:11066138
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1807)
  AUTHORS   Wu,K., Malik,K., Tian,L., Brown,D. and Miki,B.
```

```
  TITLE     Functional analysis of RPD3 histone deacetylase homologs in
            Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1807)
  AUTHORS   Wu,K., Tian,L., Malik,K., Brown,D. and Miki,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-1999) Eastern Cereal and Oilseed Research Centre,
            Agriculture and Agri-Food Canada, Ottawa, Ontario K1A 0C6, Canada
FEATURES             Location/Qualifiers
     source          1. .1807
                     /organism="Arabidopsis thaliana"
                     /cultivar="Columbia"
                     /db_xref="taxon:3702"
                     /clone="G11C3"
     gene            1. .1807
                     /gene="RPD3A"
     CDS             142. .1647
                     /gene="RPD3A"
                     /codon_start=1
                     /product="putative histone deacetylase"
                     /protein_id="AAG28474.1"
                     /db_xref="GI:11066139"
                     /translation="MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIR
                     MTHALLAHYGLLQHMQVLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRF
                     NVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCY
                     VNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGH
                     IQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSG
                     DRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALGVEVED
                     KMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQER
                     PPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIM
                     ERGKGCEVEVDESGSTKVTGVNPVGVEEEASVKMEEEGTNKGGAEQAFPPKT"
BASE COUNT      472 a    386 c    464 g    485 t
ORIGIN
```

```
  Query Match           100.0%;  Score 1807;  DB 8;  Length 1807;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1807;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 agagagcagctcccttcccctcggcgaggaggaaggaagaagaaagccagagagagagag 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 AGAGAGCAGCTCCCTTCCCCTCGGCGAGGAGGAAGGAAGAAGAAAGCCAGAGAGAGAG 60

Qy     61 agagatcatcgcagcttctcctccgaccatttgactgcgactgtgattacaacacaccgt 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGAGATCATCGCAGCTTCTCCTCCGACCATTTGACTGCGACTGTGATTACAACACACCGT 120

Qy    121 tgatcctacgaaaaagaggtaatggatactggcggcaattcgctggcgtccggacctgat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TGATCCTACGAAAAAGAGGTAATGGATACTGGCGGCAATTCGCTGGCGTCCGGACCTGAT 180

Qy    181 ggtgtgaagaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGTGTGAAGAGGAAAGTTTGTTATTTCTATGACCCTGAGGTCGGCAATTACTACTATGGC 240

Qy    241 caaggtcatcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactac 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAAGGTCATCCCATGAAGCCCCATCGCATCCGCATGACCCATGCCCTCCTCGCTCACTAC 300

Qy    301 ggtctccttcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgc 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGTCTCCTTCAGCATATGCAGGTTCTCAAGCCCTTCCCTGCCCGCGAACGTGATCTCTGC 360

Qy    361 cgcttccacgccgacgactatgtctcttttctccgcagcattacccctgaaacccagcaa 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGCTTCCACGCCGACGACTATGTCTCTTTTCTCCGCAGCATTACCCCTGAAACCCAGCAA 420

Qy    421 gatcagattcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GATCAGATTCGCCAACTTAAGCGCTTCAATGTTGGTGAAGACTGTCCCGTCTTTGACGGC 480

Qy    481 ctttattccttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaac 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTTATTCCTTTTGCCAGACCTATGCTGGAGGATCTGTTGGTGGCTCTGTCAAGCTTAAC 540

Qy    541 cacggcctctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CACGGCCTCTGCGATATTGCCATCAACTGGGCTGGTGGTCTCCATCACGCTAAGAAGTGC 600

Qy    601 gaggcctctggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaag 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GAGGCCTCTGGCTTCTGTTACGTCAATGATATCGTCTTAGCTATCCTAGAGCTCCTTAAG 660

Qy    661 cagcatgagcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggag 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CAGCATGAGCGTGTTCTTTATGTCGATATTGATATCCACCACGGGGATGGAGTGGAGGAG 720

Qy    721 gcattttatgctactgacagggttatgactgtctcgtttcataaatttggtgattacttt 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCATTTTATGCTACTGACAGGGTTATGACTGTCTCGTTTCATAAATTTGGTGATTACTTT 780

Qy    781 cccggtacaggtcacattcaggatataggttatggtagcggaaagtactattctctcaat 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CCCGGTACAGGTCACATTCAGGATATAGGTTATGGTAGCGGAAAGTACTATTCTCTCAAT 840

Qy    841 gtaccactggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatg 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GTACCACTGGATGATGGAATCGATGATGAGAGCTATCATCTGTTATTCAAGCCCATCATG 900

Qy    901 gggaaagttatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactcc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GGGAAAGTTATGGAAATTTTCCGACCAGGGGCTGTGGTATTGCAATGTGGTGCTGACTCC 960

Qy    961 ctatctggggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CTATCTGGGGATCGGTTAGGTTGCTTCAATCTTTCAATCAAAGGTCATGCTGAGTGCGTC 1020

Qy   1021 aaatttatgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AAATTTATGAGATCGTTCAATGTTCCCCTACTGCTCTTGGGTGGTGGTGGTTACACTATC 1080

Qy   1081 cgcaatgttgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CGCAATGTTGCCCGTTGCTGGTGCTACGAGACTGGAGTTGCACTTGGAGTTGAAGTTGAA 1140

Qy   1141 gacaagatgccggagcatgaatattatgaatactttggtccagactatacacttcacgtt 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACAAGATGCCGGAGCATGAATATTATGAATACTTTGGTCCAGACTATACACTTCACGTT 1200

Qy   1201 gctccaagtaacatggaaaataagaattctcgtcagatgcttgaagagattcgcaatgac 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GCTCCAAGTAACATGGAAAATAAGAATTCTCGTCAGATGCTTGAAGAGATTCGCAATGAC 1260

Qy   1261 cttctccacaatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagacca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 CTTCTCCACAATCTCTCTAAGCTTCAGCATGCTCCAAGTGTACCATTTCAGGAAAGACCA 1320

Qy   1321 cctgatacagagactcccgaggttgatgaagaccaagaagatggggataaaagatgggat 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CCTGATACAGAGACTCCCGAGGTTGATGAAGACCAAGAAGATGGGGATAAAAGATGGGAT 1380

Qy   1381 ccggattcagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaaga 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CCGGATTCAGACATGGATGTTGATGATGACCGTAAACCTATACCAAGCAGAGTAAAAAGA 1440

Qy   1441 gaagctgttgaaccagatacaaaggacaaggatggactgaaaggaattatggagcgtgga 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GAAGCTGTTGAACCAGATACAAAGGACAAGGATGGACTGAAAGGAATTATGGAGCGTGGA 1500

Qy   1501 aaaggttgtgaggtggaggtggatgagagtggaagcactaaggttacaggagtaaaccca 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 AAAGGTTGTGAGGTGGAGGTGGATGAGAGTGGAAGCACTAAGGTTACAGGAGTAAACCCA 1560

Qy   1561 gtgggagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaacaagggtggggcg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GTGGGAGTGGAGGAAGCAAGTGTGAAAATGGAAGAGGAAGGAACAAACAAGGGTGGGGCG 1620
```

```
Qy   1621 gagcaggcgtttcctcctaaaacataagactcggagcttctaatttcttgctacttttttc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GAGCAGGCGTTTCCTCCTAAAACATAAGACTCGGAGCTTCTAATTTCTTGCTACTTTTTC 1680

Qy   1681 tgtctatcaaatgttgctagttaagtttctggagttgttgttgttgtaagcactcctctg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 TGTCTATCAAATGTTGCTAGTTAAGTTTCTGGAGTTGTTGTTGTTGTAAGCACTCCTCTG 1740

Qy   1741 ttttagaggattgagcacggatatgtatttattcgttgcatgtctgaatgatgatatgat 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TTTTAGAGGATTGAGCACGGATATGTATTTATTCGTTGCATGTCTGAATGATGATATGAT 1800

Qy   1801 atgacaa 1807
          |||||||
Db   1801 ATGACAA 1807
```

```
RESULT   3
AF014824
LOCUS       AF014824                1808 bp    mRNA    linear   PLN 12-AUG-1997
DEFINITION  Arabidopsis thaliana histone deacetylase mRNA, complete cds.
ACCESSION   AF014824
VERSION     AF014824.1  GI:2318130
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1808)
  AUTHORS   Tomihama,T., Shoji,K., Hanyu,H. and Okano,T.
  TITLE     Characterization of a histone deacetylase (EST G11C3T7) in
            Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1808)
  AUTHORS   Tomihama,T., Shoji,K., Hanyu,H. and Okano,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-1997) Department of Biology, Central Research
            Institute of Electric Power Industry, 1646, Abiko, Chiba 270-11,
            Japan
FEATURES             Location/Qualifiers
     source          1. .1808
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /note="EST clone No.G11C3T7"
     CDS             143. .1648
                     /codon_start=1
                     /product="histone deacetylase"
                     /protein_id="AAB66486.1"
                     /db_xref="GI:2318131"
                     /translation="MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIR
                     MTHALLAHYGLLQHMQVLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRF
                     NVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCY
                     VNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGH
                     IQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSG
                     DRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALGVEVED
                     KMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQER
                     PPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIM
                     ERGKGCEVEVDESGSTKVTGVNPVGVEEASVKMEEEGTNKGGAEQAFPPKT"
BASE COUNT      472 a    387 c    464 g    485 t
ORIGIN
```

```
  Query Match           100.0%;  Score 1807;  DB 8;  Length 1808;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1807;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 agagagcagctcccttcccctcggcgaggaggaaggaagaagaaagccagagagagagag 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2 AGAGAGCAGCTCCCTTCCCCTCGGCGAGGAGGAAGGAAGAAGAAAGCCAGAGAGAGAGAG 61
```

```
Qy      61 agagatcatcgcagcttctcctccgaccatttgactgcgactgtgattacaacacaccgt 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      62 AGAGATCATCGCAGCTTCTCCTCCGACCATTTGACTGCGACTGTGATTACAACACACCGT 121

Qy     121 tgatcctacgaaaaagaggtaatggatactggcggcaattcgctggcgtccggacctgat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 TGATCCTACGAAAAAGAGGTAATGGATACTGGCGGCAATTCGCTGGCGTCCGGACCTGAT 181

Qy     181 ggtgtgaagaggaaagtttgttatttctatgaccctgaggtcggcaattactactatggc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 GGTGTGAAGAGGAAAGTTTGTTATTTCTATGACCCTGAGGTCGGCAATTACTACTATGGC 241

Qy     241 caaggtcatcccatgaagccccatcgcatccgcatgacccatgccctcctcgctcactac 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 CAAGGTCATCCCATGAAGCCCCATCGCATCCGCATGACCCATGCCCTCCTCGCTCACTAC 301

Qy     301 ggtctccttcagcatatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 GGTCTCCTTCAGCATATGCAGGTTCTCAAGCCCTTCCCTGCCCGCGAACGTGATCTCTGC 361

Qy     361 cgcttccacgccgacgactatgtctcttttctccgcagcattacccctgaaaacccagcaa 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CGCTTCCACGCCGACGACTATGTCTCTTTTCTCCGCAGCATTACCCCTGAAACCCAGCAA 421

Qy     421 gatcagattcgccaacttaagcgcttcaatgttggtgaagactgtcccgtctttgacggc 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 GATCAGATTCGCCAACTTAAGCGCTTCAATGTTGGTGAAGACTGTCCCGTCTTTGACGGC 481

Qy     481 ctttattccttttgccagacctatgctggaggatctgttggtggctctgtcaagcttaac 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     482 CTTTATTCCTTTTGCCAGACCTATGCTGGAGGATCTGTTGGTGGCTCTGTCAAGCTTAAC 541

Qy     541 cacggcctctgcgatattgccatcaactgggctggtggtctccatcacgctaagaagtgc 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     542 CACGGCCTCTGCGATATTGCCATCAACTGGGCTGGTGGTCTCCATCACGCTAAGAAGTGC 601

Qy     601 gaggcctctggcttctgttacgtcaatgatatcgtcttagctatcctagagctccttaag 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     602 GAGGCCTCTGGCTTCTGTTACGTCAATGATATCGTCTTAGCTATCCTAGAGCTCCTTAAG 661

Qy     661 cagcatgagcgtgttctttatgtcgatattgatatccaccacggggatggagtggaggag 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     662 CAGCATGAGCGTGTTCTTTATGTCGATATTGATATCCACCACGGGGATGGAGTGGAGGAG 721

Qy     721 gcattttatgctactgacagggttatgactgtctcgtttcataaatttggtgattacttt 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     722 GCATTTTATGCTACTGACAGGGTTATGACTGTCTCGTTTCATAAATTTGGTGATTACTTT 781

Qy     781 cccggtacaggtcacattcaggatataggttatggtagcggaaagtactattctctcaat 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     782 CCCGGTACAGGTCACATTCAGGATATAGGTTATGGTAGCGGAAAGTACTATTCTCTCAAT 841

Qy     841 gtaccactggatgatggaatcgatgatgagagctatcatctgttattcaagcccatcatg 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     842 GTACCACTGGATGATGGAATCGATGATGAGAGCTATCATCTGTTATTCAAGCCCATCATG 901

Qy     901 gggaaagttatggaaattttccgaccaggggctgtggtattgcaatgtggtgctgactcc 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     902 GGGAAAGTTATGGAAATTTTCCGACCAGGGGCTGTGGTATTGCAATGTGGTGCTGACTCC 961

Qy     961 ctatctggggatcggttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtc 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     962 CTATCTGGGGATCGGTTAGGTTGCTTCAATCTTTCAATCAAAGGTCATGCTGAGTGCGTC 1021

Qy    1021 aaatttatgagatcgttcaatgttcccctactgctcttgggtggtggtggttacactatc 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1022 AAATTTATGAGATCGTTCAATGTTCCCCTACTGCTCTTGGGTGGTGGTGGTTACACTATC 1081

Qy    1081 cgcaatgttgcccgttgctggtgctacgagactggagttgcacttggagttgaagttgaa 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1082 CGCAATGTTGCCCGTTGCTGGTGCTACGAGACTGGAGTTGCACTTGGAGTTGAAGTTGAA 1141

Qy    1141 gacaagatgccggagcatgaatattatgaatactttggtccagactatacacttcacgtt 1200
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1142 GACAAGATGCCGGAGCATGAATATTATGAATACTTTGGTCCAGACTATACACTTCACGTT 1201

Qy    1201 gctccaagtaacatggaaaataagaattctcgtcagatgcttgaagagattcgcaatgac 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1202 GCTCCAAGTAACATGGAAAATAAGAATTCTCGTCAGATGCTTGAAGAGATTCGCAATGAC 1261

Qy    1261 cttctccacaatctctctaagcttcagcatgctccaagtgtaccatttcaggaaagacca 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1262 CTTCTCCACAATCTCTCTAAGCTTCAGCATGCTCCAAGTGTACCATTTCAGGAAAGACCA 1321

Qy    1321 cctgatacagagactcccgaggttgatgaagaccaagaagatggggataaaagatgggat 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1322 CCTGATACAGAGACTCCCGAGGTTGATGAAGACCAAGAAGATGGGGATAAAAGATGGGAT 1381

Qy    1381 ccggattcagacatggatgttgatgatgaccgtaaacctataccaagcagagtaaaaaga 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1382 CCGGATTCAGACATGGATGTTGATGATGACCGTAAACCTATACCAAGCAGAGTAAAAAGA 1441

Qy    1441 gaagctgttgaaccagatacaaaggacaaggatggactgaaaggaattatggagcgtgga 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1442 GAAGCTGTTGAACCAGATACAAAGGACAAGGATGGACTGAAAGGAATTATGGAGCGTGGA 1501

Qy    1501 aaaggttgtgaggtggaggtggatgagagtggaagcactaaggttacaggagtaaaccca 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1502 AAAGGTTGTGAGGTGGAGGTGGATGAGAGTGGAAGCACTAAGGTTACAGGAGTAAACCCA 1561

Qy    1561 gtgggagtggaggaagcaagtgtgaaaatggaagaggaaggaacaaacaagggtggggcg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1562 GTGGGAGTGGAGGAAGCAAGTGTGAAAATGGAAGAGGAAGGAACAAACAAGGGTGGGGCG 1621

Qy    1621 gagcaggcgtttcctcctaaaacataagactcggagcttctaatttcttgctacttttttc 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1622 GAGCAGGCGTTTCCTCCTAAAACATAAGACTCGGAGCTTCTAATTTCTTGCTACTTTTTC 1681

Qy    1681 tgtctatcaaatgttgctagttaagtttctggagttgttgttgttgtaagcactcctctg 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1682 TGTCTATCAAATGTTGCTAGTTAAGTTTCTGGAGTTGTTGTTGTTGTAAGCACTCCTCTG 1741

Qy    1741 ttttagaggattgagcacggatatgtatttattcgttgcatgtctgaatgatgatatgat 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1742 TTTTAGAGGATTGAGCACGGATATGTATTTATTCGTTGCATGTCTGAATGATGATATGAT 1801

Qy    1801 atgacaa 1807
          |||||||
Db    1802 ATGACAA 1808
```

```
RESULT   4
ATF20D10/c
LOCUS       ATF20D10              147215 bp    DNA     linear   PLN 26-FEB-1999
DEFINITION  Arabidopsis thaliana DNA chromosome 4, BAC clone  F20D10 (ESSA
            project).
ACCESSION   AL035538
VERSION     AL035538.1  GI:4467094
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 147215)
  AUTHORS   Bevan,M., Wedler,H., Kutzner,M., Wambutt,R., Bancroft,I.,
            Mewes,H.W., Mayer,K.F.X. and Schueller,C.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 147215)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1999) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            schuelle@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosome 4 can be
            viewed at: http://websvr.mips.biochem.mpg.de/proj/thal/.
FEATURES             Location/Qualifiers
     source          1. .147215
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     gene            complement(join(830. .2566,2624. .3049))
                     /gene="F20D10.10"
     gene            830. .3049
                     /gene="F20D10.10"
     CDS             complement(join(830. .2566,2624. .3049))
                     /gene="F20D10.10"
                     /note="strong similarity to retrotransposon -like protein
                     -Arabidopsis thaliana, PID:e1248853"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB37529.1"
                     /db_xref="GI:4467095"
                     /translation="MMNGLRRTFWSSIHKKKDNNRVDDSLDRQKPTTTSRFGFFSNPS
                     TPRSETRPDSVTCSPTIPCQSWSATAISTPSPSLPASPKLQCDTSGDVTPTRNRSPLS
                     FLSVSSSSSTPSSPKSPASFSLLKSKLCFTKVRILNRNNQSSISSSKCGICLQSAKAG
                     RGTAIFTAECSHTFHFPCVASRAGDRNLLSDCPVCGASWRETSLLPLSLSSSLHESGS
                     ESDSKIRESKNNNKSLRVYNDDEPLISSPISRTGFNTIPESNEDEEEEDNDDGEFKGF
                     YVNTPSPLTTKKMLTDSVTGHVDVKLSSEAAIVAVGRGNETYSVLMKIKSPSLPTARR
                     SPVDLVTVIDVSGGNIEMVKRAMRQVISSLRETDRLSMVSFSSSSSKRLTPLRRMTANG
                     RRLARRIVDDISGDGDGMSVNDAVKKAAKVIEDRRQKNLFTTIFVLTDRNRNSAHQAQ
                     LAQPDFVTSTRFSHLEIPTHTIWLGACNHALPEDVFAKRIKSLLSLSVQDLTLNLGLV
                     SGSGQGKVTSVYSLSGRPVWLGSGLIRLGDMYGDEEREVLVELKSPSSSRSQRIMTVR
                     SRHVDPTTQEIKNYEDRALMIPRPTTVRSSDPSIARLRNLHVSTRAVAESRRLVEVND
                     YSGAERMLTSARALLVQYGLSSSDSCLRGLEAELADLNRLRGRHVAVKSPEPVVQKSE
                     PLTPTSAWRAAERLAKVAIMRKHMNRVSDLHGFENARF"
     exon            complement(830. .2566)
                     /gene="F20D10.10"
                     /number=1
     intron          complement(2567. .2623)
                     /gene="F20D10.10"
                     /number=1
     exon            complement(2624. .3049)
                     /gene="F20D10.10"
                     /number=2
     repeat_region   5063. .5098
                     /note="35bp ta tandem repeat"
     exon            complement(9755. .10264)
                     /gene="F20D10.20"
                     /number=1
     gene            9755. .12463
                     /gene="F20D10.20"
     gene            complement(join(9755. .10264,10338. .11657,11741. .11989,
                     12071. .12191,12300. .12463))
                     /gene="F20D10.20"
     CDS             complement(join(9755. .10264,10338. .11657,11741. .11989,
                     12071. .12191,12300. .12463))
                     /gene="F20D10.20"
                     /note="similarity to Caenorhabditis elegans cosmid F32B5,
                     gene F32B5.7, PID:g2088783
                     Contains Eukaryotic putative RNA-binding region RNP-1
                     signature [KGSCFLPM]
                     contains EST gb:N97112;AA728688"
                     /codon_start=1
                     /product="putative protein"
                     /protein_id="CAB37530.1"
                     /db_xref="GI:4467096"
                     /translation="MDKEKEQTLEWNEAQKIDISVDLLAAAKKHLLFLGAVDRNRCLY
                     DGPALQRAIYRYNAYWLPLLAQYTESSSICQGPLVPPLDCEWVWHCHRLNPVRYKTDC
                     EQFYGRVLDNSGVVSSVNGNCKSQTETLWKRLYPTEPYDLDFANAISEPADVSALEKC
                     TTYDLVLAVKRQSPFFYQVSRAHVDNDVFLQEAVARYKAFLYLIKGNRERSIKLFCVP
                     TYDIDLIWHTHQLHAISYCNDLTKMIGKVLEHDDTDSDRSKGKKLDTGFSGTTAQWEE
                     TFGRRYWKAGAMNRGNTPKPVTTSPYVCSGKKSIAKEEESQNVIQYPEVKVIEVILEI
                     VGVKNLPDAHKGKVFVLFSKTQPDSLFNAERRLTVLSESCGEKQVALFQCEPTGELSF
```

```
                QLMSSKSKSLGFTSLSFSEFLSPVTKLSVEKWLELTPTKRGKADDPNPISLRVAVSFT
                PPTRSPTVLHLVQARPSLKGSCFLPMLRKVRLAKSFTRVVDETETEVINLQMRNSNDA
                APKGDRRQVIGVKECGETYVLAEYDGTFWSLLDSKWSLKQTCNPATDGPLFELSGTRM
                VKVYSGRKLEYEPKHCSKLRSEQDFMTAVEFSKQHPYGKAVGLLDLKFGSIEANEKWL
                VLPGMVSSFILSDLLKKEGFSAAAKDTVKANGITEESTEIDVLSQEKLEEETMMDVDT
                TTPVAVAAEKINGGARCFSKELSGNMIEEEGGHCGGCGGCGGCGGGGGCGGGGRCGGM
                TKIEGCGGGSCTGGSTGCGNCGGGCGNMMKNNANGNAPSVENDAVTA"
intron          complement(10265. .10337)
                /gene="F20D10.20"
                /number=1
exon            complement(10338. .11657)
                /gene="F20D10.20"
                /number=2
intron          complement(11658. .11740)
                /gene="F20D10.20"
                /number=2
exon            complement(11741. .11989)
                /gene="F20D10.20"
                /number=3
intron          complement(11990. .12070)
                /gene="F20D10.20"
                /number=3
exon            complement(12071. .12191)
                /gene="F20D10.20"
                /number=4
intron          complement(12192. .12299)
                /gene="F20D10.20"
                /number=4
exon            complement(12300. .12463)
                /gene="F20D10.20"
                /number=5
gene            complement(join(13386. .14030,14118. .14738,14828. .15394,
                15491. .15569,15669. .15730,16043. .16069))
                /gene="F20D10.30"
gene            13386. .16069
                /gene="F20D10.30"
CDS             complement(join(13386. .14030,14118. .14738,14828. .15394,
                15491. .15569,15669. .15730,16043. .16069))
                /gene="F20D10.30"
                /note="strong similarity to dnaK-type molecular chaperone
                PHSP1 -Pisum sativum, PID:g20835
                Contains ATP/GTP-binding site motif A (P-loop) [AEAYLGKS],
                Heat shock hsp70 proteins family signatures [IDLGTTNS],
                Hsp70_2 [VFDLGGGTFDVSIL], Hsp70_3 [VLLVGGMTRVPKVQE]
                contains EST gb:T44452, Z24523, T04782, Z17893, Z17973"
                /codon_start=1
                /product="heat shock protein 70 like protein"
                /protein_id="CAB37531.1"
                /db_xref="GI:4467097"
                /translation="MASVSAFKSVSANGKNSMFGKLGYLARPFCSRPVGNDVIGIDLG
                TTNSCVSVMEGKTARVIENAEGSRTTPSVVAMNQKGELLVGTPAKRQAVTNPTNTIFG
                SKRLIGRRFDDPQTQKEMKMVPYKIVKAPNGDAWVEANGQKFSPSQIGANVLTKMKET
                AEAYLGKSINKAVVTVPAYFNDAQRQATKDAGKIAGLDVQRIINEPTAAALSYGMNNK
                EGVIAVFDLGGGTFDVSILEISSGVFEVKATNGDTFLGGEDFDNTLLEYLVNEFKRSD
                NIDLTKDNLALQRLREAAEKAKIELSSTTQTEINLPFITADASGAKHLNITLTRSKFE
                GLVGKLIERTRSPCQNCLKDAGVTIKEVDEVLLVGGMTRVPKVQEIVSEIFGKSPCKG
                VNPDEAVAMGAAIQGGILRGDVKDLLLLDVVPLSLGIETLGAVFTKLIPRNTTIPTKK
                SQVFSTAADNQMQVGIKVLQGEREMAADNKVLGEFDLVGIPPAPRGMPQIEVTFDIDA
                NGITTVSAKDKATGKEQNITIRSSGGLSDDEINRMVKEAELNAQKDQEKKQLIDLRNS
                ADTTIYSVEKSLSEYREKIPAEIASEIETAVSDLRTAMAGEDVEDIKAKVEAANKAVS
                KIGEHMSKGSGSSGSDGSSGEGTSGTEQTPEAEFEEASGSRK"
exon            complement(13386. .14030)
                /gene="F20D10.30"
                /number=1
intron          complement(14031. .14117)
                /gene="F20D10.30"
                /number=1
exon            complement(14118. .14738)
                /gene="F20D10.30"
                /number=2
intron          complement(14739. .14827)
                /gene="F20D10.30"
                /number=2
exon            complement(14828. .15394)
                /gene="F20D10.30"
                /number=3
intron          complement(15395. .15490)
                /gene="F20D10.30"
                /number=3
exon            complement(15491. .15569)
                /gene="F20D10.30"
                /number=4
intron          complement(15570. .15668)
                /gene="F20D10.30"
                /number=4
exon            complement(15669. .15730)
                /gene="F20D10.30"
                /number=5
intron          complement(15731. .16042)
                /gene="F20D10.30"
                /number=5
exon            complement(16043. .16069)
                /gene="F20D10.30"
                /number=6
gene            16556. .19503
                /gene="F20D10.40"
exon            complement(16556. .16790)
                /gene="F20D10.40"
                /number=1
gene            complement(join(16556. .16790,16869. .17056,17133. .17377,
                17453. .17618,17710. .17975,18152. .18358,18789. .19503))
                /gene="F20D10.40"
CDS             complement(join(16556. .16790,16869. .17056,17133. .17377,
                17453. .17618,17710. .17975,18152. .18358,18789. .19503))
```

Query Match 53.1%; Score 959; DB 8; Length 147215;
Best Local Similarity 84.3%; Pred. No. 1.1e-258;
Matches 1199; Conservative 0; Mismatches 10; Indels 214; Gaps 2;

```
Qy    135 agaggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaa 194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87082 AGAGGTAATGGATACTGGCGGCAATTCGCTGGCGTCCGGACCTGATGGTGTGAAGAGGAA 87023

Qy    195 agtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccat 254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  87022 AGTTTGTTATTTCTATGACCCTGAGGTCGGCAATTACTACTATGGCCAAGGTCATCCCAT 86963

Qy    255 gaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagca 314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86962 GAAGCCCCATCGCATCCGCATGACCCATGCCCTCCTCGCTCACTACGGTCTCCTTCAGCA 86903

Qy    315 tatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccga 374
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  86902 TATGCAGGTTCTCAAGCCCTTCCCTGCCCGCGACCGTGATCTCTGCCGCTTCCACGCCGA 86843

Qy    375 cgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgcca 434
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86842 CGACTATGTCTCTTTTCTCCGCAGCATTACCCCTGAAACCCAGCAAGATCAGATTCGCCA 86783

Qy    435 acttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttg 494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86782 ACTTAAGCGCTTCAATGTTGGTGAAGACTGTCCCGTCTTTGACGGCCTTTATTCCTTTTG 86723

Qy    495 ccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcga 554
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86722 CCAGACCTATGCTGGAGGATCTGTTGGTGGCTCTGTCAAGCTTAACCACGGCCTCTGCGA 86663

Qy    555 tattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggctt 614
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86662 TATTGCCATCAACTGGGCTGGTGGTCTCCATCACGCTAAGAAGTGCGAGGCCTCTGGCTT 86603

Qy    615 ctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatg------- 667
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86602 CTGTTACGTCAATGATATCGTCTTAGCTATCCTAGAGCTCCTTAAGCAGCATGAGGTTTG 86543

Qy    668 ---------------------------------------------------------- 667
```

```
Db  86542 TCCACCTGCACCGTTTTGATATATGCTAGGGATGGTTATGTCGTTTCTGAATCTCTAACT 86483

Qy    668 -----------------agcgtgttctttatgtcgatattgatatccaccacggggatgg 710
                          ||||||||||||||||||||||||||||||||||||||||||||
Db  86482 CTTCCTCTCTCGTTTGCAGCGTGTTCTTTATGTCGATATTGATATCCACCACGGGGATGG 86423

Qy    711 agtggaggaggcattttatgctactgacagggttatgactgtctcgtttcataaatttgg 770
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86422 AGTGGAGGAGGCATTTTATGCTACTGACAGGGTTATGACTGTCTCGTTTCATAAATTTGG 86363

Qy    771 tgattactttcccggtacaggtcacattcaggatataggttatggtagcggaaagtacta 830
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86362 TGATTACTTTCCCGGTACAGGTCACATTCAGGATATAGGTTATGGTAGCGGAAAGTACTA 86303

Qy    831 ttctctcaatgtaccactggatgatggaatcgatgatgagagctatcatctgttattcaa 890
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86302 TTCTCTCAATGTACCACTGGATGATGGAATCGATGATGAGAGCTATCATCTGTTATTCAA 86243

Qy    891 gcccatcatggggaaagttatggaaatttttccgaccaggggctgtggtattgcaatgtgg 950
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86242 GCCCATCATGGGGAAAGTTATGGAAATTTTCCGACCAGGGGCTGTGGTATTGCAATGTGG 86183

Qy    951 tgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaaaggtcatgc 1010
          |||||| ||  | ||||| ||| |||| || ||||| ||||||||||||||||||||||
Db  86182 TGCTGATTCATTGTCTGGTGATAGGTTGGGGTGCTTTAATCTTTCAATCAAAGGTCATGC 86123

Qy   1011 tgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttgggtggtggtgg 1070
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  86122 TGAGTGCGTCAAATTTATGAGATCGTTCAATGTTCCCCTACTGCTCTTGGGTGGTGGTGG 86063

Qy   1071 ttacactatccgcaatgttgcccgttgctggtgctacg---------------------- 1108
          |||||||||||||||||||||||||||||||||||
Db  86062 TTACACTATCCGCAATGTTGCCCGTTGCTGGTGCTACGAGGTACTTTCTTCTGTTCATTT 86003

Qy   1109 ------------------------------------------------------------ 1108

Db  86002 TATGCTGCTCTAGCTTTGCATTGAAATTTCCTTGGTCTCTGTACCAAGGTTAGCCCTCTG 85943

Qy   1109 ----------------------------------------------agactggagttg 1120
                                                         ||||||||||||
Db  85942 TGACGCCTCTCTTTCTTTGAATTTATTACAACAAACTGTTTATATGACAGACTGGAGTTG 85883

Qy   1121 cacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggtc 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  85882 CACTTGGAGTTGAAGTTGAAGACAAGATGCCGGAGCATGAATATTATGAATACTTTGGTC 85823

Qy   1181 cagactatacacttcacgttgctccaagtaacatggaaaatgaagaattctcgtcagatgc 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  85822 CAGACTATACACTTCACGTTGCTCCAAGTAACATGGAAAATAAGAATTCTCGTCAGATGC 85763

Qy   1241 ttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagtg 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  85762 TTGAAGAGATTCGCAATGACCTTCTCCACAATCTCTCTAAGCTTCAGCATGCTCCAAGTG 85703

Qy   1301 taccatttcaggaaagaccacctgatacagagactcccgaggt 1343
          |||||||||||||||||||||||||||||||||||||||||||
Db  85702 TACCATTTCAGGAAAGACCACCTGATACAGAGACTCCCGAGGT 85660


RESULT   5
ATCHRIV89/c
LOCUS       ATCHRIV89              199789 bp    DNA     linear   PLN 16-MAR-2000
DEFINITION  Arabidopsis thaliana DNA chromosome 4, contig fragment No. 89.
ACCESSION   AL161593
VERSION     AL161593.2  GI:7270796
KEYWORDS    .
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 162375 to 199789)
  AUTHORS   Murphy,G., Ridley,P., Hudson,S., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   2  (bases 195922 to 196568)
  AUTHORS   Robben,J., Braeken,M., Grymonprez,B., Volckaert,G., Mewes,H.W.,
            Lemcke,K. and Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 68137; 40958 to 134717; 99060 to 190997)
  AUTHORS   Wedler,H., Kutzner,M., Wambutt,R., Mewes,H.W., Lemcke,K. and
            Mayer,K.F.X.
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 199789)
  AUTHORS   EU Arabidopsis sequencing,project.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer
            Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
            lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
            Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
            Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
            E-mail: michael.bevan@bbsrc.ac.uk
COMMENT     Information on performance of analysis and a more detailed
            annotation of this entry and other sequences of chromosomes 3, 4
            and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/
            this fragment has an overlap with ATCHRIV88 at the 5' end and an
            overlap with ATCHRIV90 at the 3' end.
FEATURES             Location/Qualifiers
     source          1. .199789
                     /organism="Arabidopsis thaliana"
                     /variety="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="4"
     exon            complement(5433. .5537)
                     /gene="AT4g38130"
                     /number=1
     gene            5433. .7997
                     /gene="AT4g38130"
     gene            complement(join(5433. .5537,5612. .5683,6112. .6169,
                     6258. .6328,6584. .6814,6945. .7385,7470. .7997))
                     /gene="AT4g38130"
     CDS             complement(join(5433. .5537,5612. .5683,6112. .6169,
                     6258. .6328,6584. .6814,6945. .7385,7470. .7997))
                     /gene="AT4g38130"
                     /note="contains EST gb:R90034, AI994124.1, W43397,
                     AW004334.1, F15302, AA605428"
                     /codon_start=1
                     /product="Histone deacetylase"
                     /protein_id="CAB80478.1"
                     /db_xref="GI:7270797"
                     /translation="MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIR
                     MTHALLAHYGLLQHMQVLKPFPARDRDLCRFHADDYVSFLRSITPETQQDQIRQLKRF
                     NVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCY
                     VNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGH
                     IQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSG
                     DRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALGVEVED
                     KMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQER
                     PPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIM
                     ERGKGCEVEVDESGSTKVTGVNPVGVEEASVKMEEEGTNKGGAEQAFPPKT"
     intron          complement(5538. .5611)
                     /gene="AT4g38130"
                     /number=1
     exon            complement(5612. .5683)
                     /gene="AT4g38130"
                     /number=2
     intron          complement(5684. .6111)
                     /gene="AT4g38130"
                     /number=2
     exon            complement(6112. .6169)
                     /gene="AT4g38130"
                     /number=3
     intron          complement(6170. .6257)
                     /gene="AT4g38130"
                     /number=3
     exon            complement(6258. .6328)
```

```
                /gene="AT4g38130"
                /number=4
     intron     complement(6329. .6583)
                /gene="AT4g38130"
                /number=4
     exon       complement(6584. .6814)
                /gene="AT4g38130"
                /number=5
     intron     complement(6815. .6944)
                /gene="AT4g38130"
                /number=5
     exon       complement(6945. .7385)
                /gene="AT4g38130"
                /number=6
     intron     complement(7386. .7469)
                /gene="AT4g38130"
                /number=6
     exon       complement(7470. .7997)
                /gene="AT4g38130"
                /number=7
     gene       8808. .9245
                /gene="AT4g38140"
     exon       complement(8808. .9245)
                /gene="AT4g38140"
                /number=1
     gene       complement(8808. .9245)
                /gene="AT4g38140"
     CDS        complement(8808. .9245)
                /gene="AT4g38140"
                /note="similarity to RING-H2 zinc finger protein RHA1a
                -Arabidopsis thaliana, PID:g3790554"
                /codon_start=1
                /product="putative protein"
                /protein_id="CAB80479.1"
                /db_xref="GI:7270798"
                /translation="MISMLFPRSPLCTAAIVFYTCVCIPLGRLKKNGGDADAHDDDGY
                NLVGVMFGDKEKEEEICCPICLVEFEAEDAVTHLPRCAHLFHINCIEPWLLRGHLTCP
                LCRSFVLAPTPPTQNVNNAHSSSTLYLSIFFFFCIFLHLLGYL"
     gene       complement(10151. .11059)
                /gene="AT4g38150"
     gene       10151. .11059
                /gene="AT4g38150"
     exon       complement(10151. .11059)
                /gene="AT4g38150"
                /number=1
     CDS        complement(10151. .11059)
                /gene="AT4g38150"
                /note="similarity to leaf protein (PnC401) -Pharbitis nil,
                PID:d1026846"
                /codon_start=1
                /product="putative protein"
                /protein_id="CAB80480.1"
                /db_xref="GI:7270799"
                /translation="MVPSSKAVVFARQMAKQIRVTTPSMSATRFLSTGDNGQVDEQQN
                PPEPLPNRPLRGERSSNSHREPPARQAHNLGKSDTTLSDDGFLEQFKLGVNQDSRETP
                KPEQYPQEPLPPPEDSDEIFKKMKEGGLIPNAVAMLDGLCKDGLVQEAMKLFGLMRDK
                GTIPEVVIYTAVVEAFCKAHKIEDAKRIFRKMQNNGIAPNAFSYGVLVQGLYNCNMLD
                DAVAFCSEMLESGHSPNVPTFVELVDALCRVKGVEQAQSAIDTLNQKGFAVNVKAVKE
                FMDKRAPFPSLAWEAIFKKKPTEKPF"
     gene       11352. .12353
                /gene="AT4g38160"
     exon       11352. .12353
                /gene="AT4g38160"
                /number=1
     CDS        11352. .12353
                /gene="AT4g38160"
                /note="contains EST gb:AI992903.1, H37273, AI100213,
                T41873, AA067478, AI100538"
                /codon_start=1
                /product="putative protein"
                /protein_id="CAB80481.1"
                /db_xref="GI:7270800"
                /translation="MEVTNTSSIMWFFRDKGFDDPSIDKMLRKCKQLEKAQSDVASEN
                WDYLSNIVGIQERKLPYIVSRCPKILTLRLDERLIPMVECLSSLGRNPREVASAITKF
                PPILSHSVEEKLCPLLAFFQALGVPETQLGKMILFNPRLISYSIDTKLTVIVSFLASL
                GLDQDGMIGKVLVKNPFLMGYSVDKRLRPTTEFLKSSVGLSEDGIKSVVMNFPQLLCR
                DVNKILKPNYDYLKECGFGDSQIATMVTGYPQILIKSVKNSLQPRIRFLVQVMGRGMD
                EVASYPEFFHHGLKKKVESRFKLVKKNNIDCSLREMLDCNTKKFHEKFGFSEVTASF"
     gene       13703. .15379
                /gene="AT4g38170"
     exon       13703. .15130
                /gene="AT4g38170"
                /number=1
     CDS        join(13703. .15130,15212. .15379)
                /gene="AT4g38170"
                /note="Contains Prokaryotic membrane lipoprotein lipid
                attachment site AA73-83
                contains EST gb:N37149, T14014"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAB80482.1"
                /db_xref="GI:7270801"
                /translation="MSRVEHVLNYLKRRQLENPGFLYAIEDDCGNVFWADPTCRLNYT
                YFGDTLVFDTTYRRGKRYQVPFAAFTGFNHHGQPVLFGCALILNESESSFAWLFQTWL
                QAMSAPPPPSITVEPDRLIQVAVSRVFSQTRLRFSQPLIFEETEEKLAHVFQAHPTFE
                SEFINCVTETETAAEFEASWDSIVRRYYMEDNDWLQSIYNARQQWVRVFIRDTFYGEL
                STNEGSSILNSFFQGFVDASTTMQMLIKQYEKAIDSWREKELKADYEATNSTPVMKTP
                SPMEKQAASLYTRAAFIKFQEEFVETLAIPANIISDSGTHTTYRVAKFGEVHKGHTVS
                FDSLEVKANCSCQMFEYSGIICRHILAVFSAKNVLALPSRYLLRRWTKEAKIRGTEEQ
                PEFSNGCQESLNLCFNSLRQEATKYVEEGAKSIQIYKVAMDALDEAAKKVAAASNRTP
                GTRLPNGEAYPSEEARETANATNHPGGEKERTILELTAELERTGQRCEVYRANLLSIL
                RDMEEQKFQLSLKVQNARLSLKE"
     intron     15131. .15211
                /gene="AT4g38170"
                /number=1
     exon       15212. .15379
                /gene="AT4g38170"
                /number=2
     exon       complement(15642. .15815)
                /gene="AT4g38180"
                /number=1
     gene       15642. .18344
                /gene="AT4g38180"
     gene       complement(join(15642. .15815,16152. .18344))
                /gene="AT4g38180"
     CDS        complement(join(15642. .15815,16152. .18344))
```

Query Match 53.1%; Score 959; DB 8; Length 199789;
Best Local Similarity 84.3%; Pred. No. 1.2e-258;
Matches 1199; Conservative 0; Mismatches 10; Indels 214; Gaps 2;

```
Qy    135 agaggtaatggatactggcggcaattcgctggcgtccggacctgatggtgtgaagaggaa 194
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8004 AGAGGTAATGGATACTGGCGGCAATTCGCTGGCGTCCGGACCTGATGGTGTGAAGAGGAA 7945

Qy    195 agtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccat 254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7944 AGTTTGTTATTTCTATGACCCTGAGGTCGGCAATTACTACTATGGCCAAGGTCATCCCAT 7885

Qy    255 gaagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagca 314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7884 GAAGCCCCATCGCATCCGCATGACCCATGCCCTCCTCGCTCACTACGGTCTCCTTCAGCA 7825

Qy    315 tatgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccga 374
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   7824 TATGCAGGTTCTCAAGCCCTTCCCTGCCCGCGACCGTGATCTCTGCCGCTTCCACGCCGA 7765

Qy    375 cgactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgcca 434
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7764 CGACTATGTCTCTTTTCTCCGCAGCATTACCCCTGAAACCCAGCAAGATCAGATTCGCCA 7705

Qy    435 acttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttg 494
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7704 ACTTAAGCGCTTCAATGTTGGTGAAGACTGTCCCGTCTTTGACGGCCTTTATTCCTTTTG 7645

Qy    495 ccagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcga 554
```

```
Db   7644 CCAGACCTATGCTGGAGGATCTGTTGGTGGCTCTGTCAAGCTTAACCACGGCCTCTGCGA 7585

Qy    555 tattgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggctt 614
Db   7584 TATTGCCATCAACTGGGCTGGTGGTCTCCATCACGCTAAGAAGTGCGAGGCCTCTGGCTT 7525

Qy    615 ctgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatg------- 667
Db   7524 CTGTTACGTCAATGATATCGTCTTAGCTATCCTAGAGCTCCTTAAGCAGCATGAGGTTTG 7465

Qy    668 ---------------------------------------------------------- 667
Db   7464 TCCACCTGCACCGTTTTGATATATGCTAGGGATGGTTATGTCGTTTCTGAATCTCTAACT 7405

Qy    668 -----------------agcgtgttctttatgtcgatattgatatccaccacggggatgg 710
Db   7404 CTTCCTCTCTCGTTTGCAGCGTGTTCTTTATGTCGATATTGATATCCACCACGGGGATGG 7345

Qy    711 agtggaggaggcattttatgctactgacagggttatgactgtctcgtttcataaatttgg 770
Db   7344 AGTGGAGGAGGCATTTTATGCTACTGACAGGGTTATGACTGTCTCGTTTCATAAATTTGG 7285

Qy    771 tgattactttcccggtacaggtcacattcaggatataggttatggtagcggaaagtacta 830
Db   7284 TGATTACTTTCCCGGTACAGGTCACATTCAGGATATAGGTTATGGTAGCGGAAAGTACTA 7225

Qy    831 ttctctcaatgtaccactggatgatggaatcgatgatgagagctatcatctgttattcaa 890
Db   7224 TTCTCTCAATGTACCACTGGATGATGGAATCGATGATGAGAGCTATCATCTGTTATTCAA 7165

Qy    891 gcccatcatggggaaagttatggaaatttccgaccaggggctgtggtattgcaatgtgg 950
Db   7164 GCCCATCATGGGGAAAGTTATGGAAATTTTCCGACCAGGGGCTGTGGTATTGCAATGTGG 7105

Qy    951 tgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaaaggtcatgc 1010
Db   7104 TGCTGATTCATTGTCTGGTGATAGGTTGGGGTGCTTTAATCTTTCAATCAAAGGTCATGC 7045

Qy   1011 tgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttgggtggtggtgg 1070
Db   7044 TGAGTGCGTCAAATTTATGAGATCGTTCAATGTTCCCCTACTGCTCTTGGGTGGTGGTGG 6985

Qy   1071 ttacactatccgcaatgttgcccgttgctggtgctacg---------------------- 1108
Db   6984 TTACACTATCCGCAATGTTGCCCGTTGCTGGTGCTACGAGGTACTTTCTTCTGTTCATTT 6925

Qy   1109 ---------------------------------------------------------- 1108
Db   6924 TATGCTGCTCTAGCTTTGCATTGAAATTTCCTTGGTCTCTGTACCAAGGTTAGCCCTCTG 6865

Qy   1109 ----------------------------------------------agactggagttg 1120
Db   6864 TGACGCCTCTCTTTCTTTGAATTTATTACAACAAACTGTTTATATGACAGACTGGAGTTG 6805

Qy   1121 cacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggtc 1180
Db   6804 CACTTGGAGTTGAAGTTGAAGACAAGATGCCGGAGCATGAATATTATGAATACTTTGGTC 6745

Qy   1181 cagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatgc 1240
Db   6744 CAGACTATACACTTCACGTTGCTCCAAGTAACATGGAAAATAAGAATTCTCGTCAGATGC 6685

Qy   1241 ttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagtg 1300
Db   6684 TTGAAGAGATTCGCAATGACCTTCTCCACAATCTCTCTAAGCTTCAGCATGCTCCAAGTG 6625

Qy   1301 taccatttcaggaaagaccacctgatacagagactcccgaggt 1343
Db   6624 TACCATTTCAGGAAAGACCACCTGATACAGAGACTCCCGAGGT 6582

RESULT   6
```

```
AF282858
LOCUS       AF282858               1839 bp    mRNA    linear   PLN 11-JUL-2000
DEFINITION  Mesembryanthemum crystallinum histone deacetylase (Hdeac1) mRNA,
            complete cds.
ACCESSION   AF282858
VERSION     AF282858.1  GI:9022438
KEYWORDS    .
SOURCE      common ice plant.
  ORGANISM  Mesembryanthemum crystallinum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE   1  (bases 1 to 1839)
  AUTHORS   Scharte,J. and Baur,B.
  TITLE     Molecular cloning of histone deacetylase from Mesembryanthemum
            crystallinum
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1839)
  AUTHORS   Scharte,J. and Baur,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JUN-2000) Institut fuer Botanik, Universitaet
            Muenster, Schlossgarten 3, Muenster 48143, Germany
FEATURES             Location/Qualifiers
     source          1. .1839
                     /organism="Mesembryanthemum crystallinum"
                     /db_xref="taxon:3544"
                     /tissue_type="leaf"
     gene            1. .1839
                     /gene="Hdeac1"
     CDS             174. .1676
                     /gene="Hdeac1"
                     /codon_start=1
                     /product="histone deacetylase"
                     /protein_id="AAF82385.1"
                     /db_xref="GI:9022439"
                     /translation="MDTGGNSLPSGPDGVKRKVSYFYDPEVGNYYYGQGHPMKPHRIR
                     MTHALLAHYGLLQHMQVLKPVPARDRDLCRFHADDYVSFLRGITPEMQQDQLRQLKRF
                     NVGEDCPVFDGLYSFCQTYAGGSVGGAVKLNHGLCDIAINWAGGLHHAKKSEASGFCY
                     VNDIVLAILELLKVHERVLYVDIDIHHGDGVEEAFYTTERVMTVSFHKFGDYFPGTET
                     FVYIGYGKGKYYSLNVPLDDGIDDESYQSLFKPIMGKVMEIFKPGAVVLQCGADSLSG
                     DRLGCFNLSIKGHAECVKYMISFNVPLLLVGGGGYTIRYVARCWCYETGVALGFELDD
                     KIPQHEYYEYFGPDYTLHVAPSNMENKNSRPMLDDIRGKLLEYLSSLQHAPSVQFQER
                     PPEAEIPEEDEDQHDPEGRWDPDSDMEVDDRSISIRSRVKREIMEPDLKDRDDVKDAV
                     DHPRDMDVSYGAGNTSVKAPTTSSDQPEGCVEMEQDNSNRPPDQGPNMNS"
BASE COUNT      468 a    383 c    452 g    536 t
ORIGIN
```

```
  Query Match           44.0%;  Score 795;  DB 8;  Length 1839;
  Best Local Similarity  72.1%;  Pred. No. 7.8e-213;
  Matches 1035;  Conservative    0;  Mismatches  400;  Indels    0;  Gaps    0;

Qy      91 ttgactgcgactgtgattacaacacaccgttgatcctacgaaaaagaggtaatggatact 150
Db     123 TTTTCTCCGGCGACCTTCATTCTTCTCCGGCTAGCCAGCCACTGTTTGGTGATGGATACC 182

Qy     151 ggcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctat 210
Db     183 GGAGGCAATTCACTTCCATCTGGACCAGATGGGGTAAAGCGTAAAGTCTCGTACTTTTAT 242

Qy     211 gaccctgaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatc 270
Db     243 GATCCAGAAGTTGGAAACTATTACTATGGGCAAGGTCACCCTATGAAGCCACATCGAATT 302

Qy     271 cgcatgacccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaag 330
Db     303 CGTATGACACATGCTCTTCTAGCTCACTATGGGCTCCTTCAACATATGCAAGTCCTGAAG 362

Qy     331 cccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctctttt 390
Db     363 CCAGTCCCTGCAAGAGATAGAGACCTTTGCAGGTTCCACGCTGATGATTACGTGTCCTTC 422

Qy     391 ctccgcagcattacccctgaaaacccagcaagatcagattcgccaacttaagcgcttcaat 450
```

```
Db    423 CTCAGGGGAATTACACCAGAAATGCAACAGGATCAGTTAAGGCAATTGAAGAGGTTTAAT 482

Qy    451 gttggtgaagactgtcccgtctttgacggcctttattccttttgccagacctatgctgga 510
Db    483 GTTGGTGAAGATTGCCCTGTATTTGATGGCCTTTACTCCTTTTGTCAGACATATGCTGGA 542

Qy    511 ggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgg 570
Db    543 GGTTCAGTTGGTGGTGCTGTAAAATTGAATCATGGGCTTTGTGACATTGCAATTAACTGG 602

Qy    571 gctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgat 630
Db    603 GCAGGTGGTTTGCATCATGCAAAAAAAAGTGAGGCGTCGGGTTTCTGCTATGTGAATGAT 662

Qy    631 atcgtcttagctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatatt 690
Db    663 ATTGTATTGGCAATTTTGGAACTTCTCAAAGTGCATGAGCGTGTTCTGTATGTGGACATT 722

Qy    691 gatatccaccacggggatggagtggaggaggcattttatgctactgacagggttatgact 750
Db    723 GATATTCACCATGGTGATGGTGTCGAAGAGGCCTTTTATACTACAGAAAGAGTCATGACA 782

Qy    751 gtctcgtttcataaatttggtgattactttcccggtacaggtcacattcaggatataggt 810
Db    783 GTGTCTTTTCACAAGTTTGGAGATTATTTTCCTGGCACAGAGACATTCGTGTATATTGGA 842

Qy    811 tatggtagcggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgaag 870
Db    843 TATGGAAAGGGAAAGTATTATTCTCTTAATGTGCCCCTAGATGATGGCATTGATGATGAG 902

Qy    871 agctatcatctgttattcaagcccatcatggggaaagttatggaaatttccgaccaggg 930
Db    903 AGCTACCAGTCCTTGTTTAAACCCATTATGGGAAAAGTGATGGAGATTTTTAAGCCTGGG 962

Qy    931 gctgtggtattgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaat 990
Db    963 GCTGTTGTCCTACAATGTGGGGCTGATTCATTGTCTGGTGACAGGTTGGGCTGTTTTAAC 1022

Qy    991 ctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttccccta 1050
Db   1023 CTCTCCATCAAGGGTCATGCAGAGTGTGTCAAATATATGATATCCTTCAATGTCCCACTC 1082

Qy   1051 ctgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgag 1110
Db   1083 CTGTTGGTGGGCGGAGGTGGCTATACTATAAGATATGTAGCCCGTTGCTGGTGTTATGAG 1142

Qy   1111 actggagttgcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaa 1170
Db   1143 ACTGGGGTTGCCCTTGGGTTTGAGCTTGATGATAAGATCCCTCAACATGAATACTATGAG 1202

Qy   1171 tactttggtccagactatacacttcacgttgctccaagtaacatggaaaatgaagaattct 1230
Db   1203 TATTTCGGTCCTGACTATACTCTCCACGTTGCTCCGAGTAATATGGAGAACAAGAACTCC 1262

Qy   1231 cgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcat 1290
Db   1263 CGACCAATGCTGGATGATATTAGAGGAAAGCTTCTTGAGTACCTCTCAAGTCTACAGCAT 1322

Qy   1291 gctccaagtgtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaa 1350
Db   1323 GCACCAAGTGTGCAATTCCAAGAACGTCCACCTGAAGCTGAGATTCCAGAGGAAGATGAA 1382

Qy   1351 gaccaagaagatggggataaaagatgggatccggattcagacatggatgttgatgatgac 1410
Db   1383 GATCAGCACGATCCAGAGGGAAGGTGGGATCCAGATTCTGATATGGAGGTTGATGACCGT 1442

Qy   1411 cgtaaacctataccaagcagagtaaaaagagaagctgttgaaccagatacaaaggacaag 1470
Db   1443 TCCATTTCTATCCGAAGCAGAGTGAAGAGAGAGATTATGGAGCCTGACCTTAAAGATCGA 1502

Qy   1471 gatggactgaaaggaattatggagcgtggaaaaggttgtgaggtggaggtggatg 1525
```

```
Db   1503 GACGATGTCAAGGATGCTGTGGATCATCCAAGAGACATGGATGTGTCGTATGGTG 1557


RESULT   7
AF384032
LOCUS       AF384032                1906 bp    mRNA    linear   PLN 26-JUN-2001
DEFINITION  Zea mays histone deacetylase HDA101 (hda101) mRNA, complete cds.
ACCESSION   AF384032
VERSION     AF384032.1  GI:14550103
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1906)
  AUTHORS   Chandler,V.L., Kaeppler,S.M., Kaeppler,H.F. and Cone,K.C.
  TITLE     Sequences from the Plant Chromatin Consortium (NSF Plant Genome
            project 9975930)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1906)
  AUTHORS   Bergstrom,D., Springer,N.M., Schmitt,L.T., Guthrie,E.,
            Sidorenko,L., Selinger,D., Kaeppler,S.M. and Cone,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-MAY-2001) Biological Sciences, Univ. of Missouri, 101
            Tucker Hall, Columbia, MO 65203, USA
FEATURES             Location/Qualifiers
     source          1. .1906
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
     gene            1. .1906
                     /gene="hda101"
     CDS             38. .1591
                     /gene="hda101"
                     /note="similar to mammalian HDAC1"
                     /codon_start=1
                     /product="histone deacetylase HDA101"
                     /protein_id="AAK67142.1"
                     /db_xref="GI:14550104"
                     /translation="MDPSSAGSGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPM
                     KPHRIRMTHSLLARYGLLNQMQVYRPNPARDRDLCRFHADDYINFLRSVTPETQQDQI
                     RLLKRFNVGEDCPVFDGLYSFCQTYAGASVGGAVKLNHGHDIAINWSGGLHHAKKCEA
                     SGFCYVNDIVLAILELLKHHERVLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYF
                     PGTGDIRDIGHSKGKYYSLNVPLDDGIDDESYQSLFKPIMGKVMEVFRPGAVVLQCGA
                     DSLSGDRLGCFNLSIKGHAECVRYMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALG
                     QEPEDKMPVNEYYEYFGPDYTLHVAPSNMENKNTRQQLDDIRSKLLDNLSKLRHAPSV
                     HFQERVPDTEIPEQDEDQDDPDERHDPDSDMEVDDHKAVEESSRRSILGIKIKREFGE
                     NATRVQDGGRVASEHRGLEPMAEDIGSSKQAPQADASAMAIDEPSNVKNEPESSTKLQ
                     GQAAAYHKP"
BASE COUNT      488 a    492 c    492 g    434 t
ORIGIN


  Query Match             41.0%;  Score 740.4;  DB 8;  Length 1906;
  Best Local Similarity   75.0%;  Pred. No. 2e-197;
  Matches  953;  Conservative    0;  Mismatches  311;  Indels    6;  Gaps    2;

Qy    144 ggatactggcggcaattcgctggcgtc---cggacctgatggtgtgaagaggaaagtttg 200
Db     55 GGGCTCCGGCGGCAACTCCCTCCCGTCCGTCGGCCCCGACGGGCAGAAGCGGCGCGTGTG 114

Qy    201 ttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaagcc 260
Db    115 CTACTTCTACGACCCGGATGTGGGCAACTACTACTACGGGCAGGGCCATCCGATGAAGCC 174

Qy    261 ccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatatgca 320
Db    175 GCACCGCATCCGGATGACGCACTCGCTGCTGGCGCGCTACGGCCTCCTCAACCAGATGCA 234

Qy    321 ggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgacta 380
Db    235 GGTGTACCGCCCCAACCCGGCCCGCGACCGCGACCTCTGCCGCTTCCACGCCGACGACTA 294
```

```
Qy    381 tgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaacttaa 440
Db    295 CATCAACTTCCTGCGCTCCGTCACGCCGGAAACGCAGCAGGACCAGATCCGCCTGCTCAA 354

Qy    441 gcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagac 500
Db    355 GCGCTTCAACGTCGGCGAGGACTGCCCCGTCTTCGACGGCCTCTACAGCTTCTGCCAGAC 414

Qy    501 ctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgc 560
Db    415 CTATGCGGGCGCCTCCGTCGGCGGGGCCGTCAAGCTCAACCACGGCC---ATGACATCGC 471

Qy    561 catcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgtta 620
Db    472 AATCAACTGGTCGGGGGGCCTGCACCACGCCAAGAAGTGCGAGGCGTCGGGCTTCTGCTA 531

Qy    621 cgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttcttta 680
Db    532 CGTCAATGACATCGTGCTCGCCATACTCGAGCTGCTCAAGCATCACGAGAGAGTTCTGTA 591

Qy    681 tgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactgacag 740
Db    592 TGTCGATATCGATATCCACCATGGTGATGGAGTGGAGGAGGCTTTCTACACAACAGATAG 651

Qy    741 ggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattca 800
Db    652 GGTTATGACTGTCTCGTTCCACAAGTTTGGTGATTATTTCCCAGGAACAGGGGATATCCG 711

Qy    801 ggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaat 860
Db    712 TGACATTGGGCACTCAAAAGGGAAGTACTACTCCCTGAATGTCCCTCTAGATGATGGGAT 771

Qy    861 cgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaatttt 920
Db    772 TGATGATGAAAGCTACCAGTCCCTTTTTAAGCCAATCATGGGCAAAGTTATGGAGGTTTT 831

Qy    921 ccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttagg 980
Db    832 CCGCCCTGGTGCAGTTGTGCTTCAGTGTGGTGCTGATTCCTTGTCTGGGGATAGGTTGGG 891

Qy    981 ttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaa 1040
Db    892 CTGCTTCAACCTCTCAATCAAAGGTCATGCGGAATGTGTTAGGTATATGAGGTCTTTCAA 951

Qy   1041 tgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctg 1100
Db    952 CGTTCCATTGTTGCTCCTTGGTGGTGGTGGATATACCATAAGAAATGTTGCACGCTGTTG 1011

Qy   1101 gtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatga 1160
Db   1012 GTGTTATGAGACTGGAGTTGCTCTTGGCCAAGAGCCTGAAGACAAGATGCCTGTTAATGA 1071

Qy   1161 atattatgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaa 1220
Db   1072 GTACTATGAATACTTCGGTCCAGATTACACTCTTCATGTTGCACCAAGTAACATGGAGAA 1131

Qy   1221 taagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaa 1280
Db   1132 CAAAAATACACGACAACAACTGGATGATATACGATCTAAACTTCTGGATAATCTTTCAAA 1191

Qy   1281 gcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcccga 1340
Db   1192 ACTTCGACATGCTCCTAGTGTCCACTTTCAAGAGAGAGTTCCTGACACAGAAATACCTGA 1251

Qy   1341 ggttgatgaagaccaagaagatggggataaaagatgggatccggattcagacatggatgt 1400
Db   1252 GCAAGATGAAGATCAAGATGATCCAGATGAACGACATGATCCTGACTCTGATATGGAAGT 1311

Qy   1401 tgatgatgac 1410
Db   1312 GGATGACCAC 1321
```

```
RESULT   8
AR168366
LOCUS       AR168366               1943 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 7 from patent US 6287843.
ACCESSION   AR168366
VERSION     AR168366.1  GI:17904244
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1943)
  AUTHORS   Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
  TITLE     Maize histone deacetylases and their use
  JOURNAL   Patent: US 6287843-A 7 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .1943
                     /organism="unknown"
BASE COUNT      510 a    498 c    497 g    438 t
ORIGIN
```

Query Match 41.0%; Score 740.4; DB 6; Length 1943;
Best Local Similarity 75.0%; Pred. No. 2e-197;
Matches 953; Conservative 0; Mismatches 311; Indels 6; Gaps 2;

```
Qy    144 ggatactggcggcaattcgctggcgtc---cggacctgatggtgtgaagaggaaagtttg 200
Db     74 GGGCTCCGGCGGCAACTCCCTCCCGTCCGTCGGCCCCGACGGGCAGAAGCGGCGCGTGTG 133

Qy    201 ttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaagcc 260
Db    134 CTACTTCTACGACCCGGATGTGGGCAACTACTACTACGGGCAGGGCCATCCGATGAAGCC 193

Qy    261 ccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatatgca 320
Db    194 GCACCGCATCCGGATGACGCACTCGCTGCTGGCGCGCTACGGCCTCCTCAACCAGATGCA 253

Qy    321 ggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgacta 380
Db    254 GGTGTACCGCCCCAACCCGGCCCGCGACCGCGACCTCTGCCGCTTCCACGCCGACGACTA 313

Qy    381 tgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaacttaa 440
Db    314 CATCAACTTCCTGCGCTCCGTCACGCCGGAAACGCAGCAGGACCAGATCCGCCTGCTCAA 373

Qy    441 gcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagac 500
Db    374 GCGCTTCAACGTCGGCGAGGACTGCCCCGTCTTCGACGGCCTCTACAGCTTCTGCCAGAC 433

Qy    501 ctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgc 560
Db    434 CTATGCGGGCGCCTCCGTCGGCGGGGCCGTCAAGCTCAACCACGGCC---ATGACATCGC 490

Qy    561 catcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgtta 620
Db    491 AATCAACTGGTCGGGGGGCCTGCACCACGCCAAGAAGTGCGAGGCGTCGGGCTTCTGCTA 550

Qy    621 cgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttcttta 680
Db    551 CGTCAATGACATCGTGCTCGCCATACTCGAGCTGCTCAAGCATCACGAGAGAGTTCTGTA 610

Qy    681 tgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactgacag 740
Db    611 TGTCGATATCGATATCCACCATGGTGATGGAGTGGAGGAGGCTTTCTACACAACAGATAG 670

Qy    741 ggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattca 800
Db    671 GGTTATGACTGTCTCGTTCCACAAGTTTGGTGATTATTTCCCAGGAACAGGGGATATCCG 730

Qy    801 ggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaat 860
Db    731 TGACATTGGGCACTCAAAAGGGAAGTACTACTCCCTGAATGTCCCTCTAGATGATGGGAT 790
```

```
Qy    861 cgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaatttt 920
          ||||||||| ||||| ||      | || ||||| |||||||| |||||||||||  ||||
Db    791 TGATGATGAAAGCTACCAGTCCCTTTTTAAGCCAATCATGGGCAAAGTTATGGAGGTTTT 850

Qy    921 ccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttagg 980
          ||| || || || || ||   | || ||||||||||| ||| | |||||||||  |||| ||
Db    851 CCGCCCTGGTGCAGTTGTGCTTCAGTGTGGTGCTGATTCCTTGTCTGGGGATAGGTTGGG 910

Qy    981 ttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaa 1040
          ||||||||  || ||||||||||||||||| || || ||  |  | ||||||  || |||||
Db    911 CTGCTTCAACCTCTCAATCAAAGGTCATGCGGAATGTGTTAGGTATATGAGGTCTTTCAA 970

Qy   1041 tgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctg 1100
          |||||  |  ||||| | |||||||||||| || || ||  | |||||||| || || ||
Db    971 CGTTCCATTGTTGCTCCTTGGTGGTGGTGGATATACCATAAGAAATGTTGCACGCTGTTG 1030

Qy   1101 gtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatga 1160
          ||| || |||||||||||||| |||||    ||   |||||||||||||| |   ||||
Db   1031 GTGTTATGAGACTGGAGTTGCTCTTGGCCAAGAGCCTGAAGACAAGATGCCTGTTAATGA 1090

Qy   1161 atattatgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaaa 1220
          || |||||||||||| |||||||| || || ||||| ||||| ||||||||||||||| ||
Db   1091 GTACTATGAATACTTCGGTCCAGATTACACTCTTCATGTTGCACCAAGTAACATGGAGAA 1150

Qy   1221 taagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaa 1280
          || ||| | || ||    || || || || ||   | | |||||  | ||||| || ||
Db   1151 CAAAAATACACGACAACAACTGGATGATATACGATCTAAACTTCTGGATAATCTTTCAAA 1210

Qy   1281 gcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcccga 1340
          ||||  |||||||| ||||| |  ||||| || |||   ||||| |||| |  || ||
Db   1211 ACTTCGACATGCTCCTAGTGTCCACTTTCAAGAGAGAGTTCCTGACACAGAAATACCTGA 1270

Qy   1341 ggttgatgaagaccaagaagatggggataaaagatgggatccggattcagacatggatgt 1400
          |   |||||||| ||||| |||   ||| || ||   ||||| || || || ||||| ||
Db   1271 GCAAGATGAAGATCAAGATGATCCAGATGAACGACATGATCCTGACTCTGATATGGAAGT 1330

Qy   1401 tgatgatgac 1410
          |||||  ||
Db   1331 GGATGACCAC 1340
```

```
RESULT   9
AF332875
LOCUS       AF332875                1579 bp    mRNA    linear   PLN 07-FEB-2001
DEFINITION  Oryza sativa histone deacetylase HD1 mRNA, complete cds.
ACCESSION   AF332875
VERSION     AF332875.1  GI:12698879
KEYWORDS    .
SOURCE      Oryza sativa.
  ORGANISM  Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 1579)
  AUTHORS   Song,F. and Goodman,R.M.
  TITLE     Molecular characterization of a rice histone deacetylase gene OsHD1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1579)
  AUTHORS   Song,F. and Goodman,R.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-2000) Department of Plant Pathology, University
            of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 5706, USA
FEATURES             Location/Qualifiers
     source          1. .1579
                     /organism="Oryza sativa"
                     /cultivar="Yuanfengzao"
                     /db_xref="taxon:4530"
                     /dev_stage="seedling"
     CDS             13. .1494
                     /codon_start=1
                     /product="histone deacetylase HD1"
                     /protein_id="AAK01712.1"
                     /db_xref="GI:12698880"
                     /translation="MDASAGGGGNSLPTAGADGAKRRVCHFYDAEVGNYYCGQGHPMK
                     PHRIRMTHALLAHYGLLDQMQVLKPHPARDRDLCRFHADDYVAFLRSVTPETQQDQIR
                     ALKRFNVGEDCPVFDGLYSFCQTYAGGSVGGAVKLNHGHDIAINWAGGLHHAKKCEAS
                     GFCYVNDIVLAILELLKYHQRVLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFP
                     GTGDIRDIGYSEGKYYCLNVPLDDGIDDDSYQSIFKPIISKVMEMYRPGAVVLQCGAD
                     SLSGDRLGCFNLSGKGHAECVKFMRSFNVPLLLLGGGGYAIRNVARCWCYETGVALGH
                     ELTDKMPPNEYFEYFGPEYSLFVAASNMENRNTNKQLEEIKCNILDNLSKLQHAPSVQ
                     FEERIPETKLPEPDEDQEDPDERHDPDSDMVLDDHKPMGHSARSLIRNIEVKREITES
                     EAKDQHGKRLTTEHKGPEPMADDPGSSKQAPVSRRLLYPSANP"
BASE COUNT      399 a    425 c    420 g    335 t
ORIGIN
```

```
  Query Match           39.5%;  Score 714.6;  DB 8;  Length 1579;
  Best Local Similarity  73.8%;  Pred. No. 3.7e-190;
  Matches  922;  Conservative    0;  Mismatches  324;  Indels    3;  Gaps    1;

Qy    162 gctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgaggt 221
          || | || | ||  | || || |  ||| ||   || ||  | ||||| ||| | |||||
Db     48 GCCGACGGCGGGGGCCGACGGGGCCAAGCGGCGGGTGTGCCACTTCTACGACGCGGAGGT 107

Qy    222 cggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgaccca 281
          || || |||||||  || || || || || ||||||||||| |||||||| ||||||||
Db    108 GGGGAACTACTACTGCGGGCAGGGGCACCCGATGAAGCCCCACCGCATCCGGATGACCCA 167

Qy    282 tgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagccctttcctgc 341
          || || ||||| |||||||| |||||  | || |||||||| |||||||||  ||| ||
Db    168 CGCGCTGCTCGCCCACTACGGCCTCCTCGACCAGATGCAGGTGCTCAAGCCCCACCCGGC 227

Qy    342 ccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagcat 401
          ||||| || || ||||||||||||||||||||||||||||| ||| | || ||||||  | |
Db    228 GCGCGACCGCGACCTCTGCCGCTTCCACGCCGACGACTACGTCGCCTTCCTCCGCTCCGT 287

Qy    402 taccccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaaga 461
          || || || |||||||| || ||||| ||    || ||||||||||| || || || ||
Db    288 CACGCCGGAGACCCAGCAGGACCAGATCCGGGCGCTCAAGCGCTTCAACGTCGGCGAGGA 347

Qy    462 ctgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgttgg 521
          ||| |||||||| |||||||| ||   ||| |||||||||| || || ||||| || ||
Db    348 CTGCCCCGTCTTCGACGGCCTCTACAGCTTCTGCCAGACCTACGCCGGGGGATCCGTCGG 407

Qy    522 tggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggtct 581
          ||| | |||||||| ||||||||||    ||| || ||||||||||||| || || ||
Db    408 CGGCGCCGTCAAGCTCAACCACGGCC---ACGACATCGCCATCAACTGGGCCGGCGGCCT 464

Qy    582 ccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtcttagc 641
          ||| ||||| |||||||||||||||| |||||||| |||||||| || |||||| | ||
Db    465 CCACCACGCCAAGAAGTGCGAGGCCTCCGGCTTCTGCTACGTCAACGACATCGTCCTCGC 524

Qy    642 tatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccacca 701
          ||||| |||||||| ||  | || |||||||||| ||||| || || ||||||||||
Db    525 CATCCTCGAGCTCCTCAAATACCACCAGCGTGTTCTCTATGTGGACATCGATATCCACCA 584

Qy    702 cggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgtttca 761
          |||||||| |||||||||| |||||  | || |||||||| ||||| ||||||||  ||
Db    585 TGGGGATGGTGTGGAGGAGGCGTTTTACACGACGGACAGGGTGATGACGGTCTCGTTCCA 644

Qy    762 taaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagcgg 821
          || ||||| ||||| || || || |||||  | || |  || || || |||      ||
Db    645 CAAGTTTGGGGATTATTTCCCGGGAACAGGGGATATCCGTGACATTGGGTATTCAGAGGG 704

Qy    822 aaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcatct 881
          ||||| || |  || |||| || |||||||||| || |||||||  ||||| || ||
Db    705 GAAGTATTACTGCCTGAATGTCCCGCTGGATGATGGGATTGATGATGACAGCTACCAGTC 764

Qy    882 gttattcaagcccatcatggggaaagttatggaaatttttccgaccaggggctgtggtatt 941
          | |||||||| |||||  | ||||| ||||| || |  || || || || || ||  |
Db    765 CATCTTCAAGCCGATCATCAGCAAAGTCATGGAGATGTATCGTCCTGGTGCAGTCGTGCT 824

Qy    942 gcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaa 1001
          || || || ||||| ||  | || || ||| |||| || || |||||||| |||   ||
Db    825 TCAGTGCGGCGCTGATTCGTTGTCCGGCGATAGGTTGGGCTGTTTCAATCTCTCAGGGAA 884
```

```
Qy   1002 aggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttggg 1061
          ||||||||| || || || || || ||||| || ||||||||||  | || ||  | ||
Db    885 AGGTCATGCAGAATGTGTAAAGTTCATGAGGTCTTTCAATGTTCCGTTGCTTCTTCTTGG 944

Qy   1062 tggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagttgc 1121
          ||||||||| ||  | ||  | |||||||| || || |||||||| ||||| ||||||||
Db    945 TGGTGGTGGATATGCCATAAGAAATGTTGCACGCTGTTGGTGCTATGAGACAGGAGTTGC 1004

Qy   1122 acttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggtcc 1181
          ||||||   |||  |    ||||||||||     |||| |||| ||| || |||||||||
Db   1005 ACTTGGTCATGAGCTCACCGACAAGATGCCTCCAAATGAGTATTTTGAGTATTTTGGTCC 1064

Qy   1182 agactatacacttcacgttgctccaagtaacatggaaaataagaatttctcgtcagatgct 1241
          ||| || |  |||  |||||| ||||||||||||| || |  ||| |     ||   |
Db   1065 AGAATACAGTCTTTTCGTTGCAGCAAGTAACATGGAGAACAGAAATACCAACAAGCAATT 1124

Qy   1242 tgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagtgt 1301
           || || ||        || ||||  ||||||| || || ||||| ||||||||| |||||
Db   1125 GGAGGAAATAAAATGCAACATTCTGGACAATCTTTCAAAACTTCAACATGCTCCTAGTGT 1184

Qy   1302 accatttcaggaaagaccacctgatacagagactcccgaggttgatgaagaccaagaaga 1361
          | |||| | ||  ||   ||||| ||| ||   || |||   |||||||| ||||| ||
Db   1185 CCAATTTGAAGAGCGAATTCCTGAAACAAAGTTACCTGAGCCAGATGAAGATCAAGAGGA 1244

Qy   1362 tggggataaaagatgggatccggattcagacatggatgttgatgatgac 1410
          |   ||| ||||    || || || || || ||||   | ||||||
Db   1245 TCCAGATGAAAGGCACGACCCTGACTCTGATATGGTATTGGATGATCAC 1293


RESULT  10
AF035815
LOCUS       AF035815                2047 bp    mRNA    linear   PLN 26-JUN-1998
DEFINITION  Zea mays putative histone deacetylase RPD3 mRNA, complete cds.
ACCESSION   AF035815
VERSION     AF035815.1  GI:2665839
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 2047)
  AUTHORS   Rossi,V., Hartings,H. and Motto,M.
  TITLE     Identification and characterisation of an RPD3 homologue from maize
            (Zea mays L.) that is able to complement an rpd3 null mutant of
            Saccharomyces cerevisiae
  JOURNAL   Mol. Gen. Genet. 258 (3), 288-296 (1998)
  MEDLINE   98307342
REFERENCE   2  (bases 1 to 2047)
  AUTHORS   Rossi,V., Hartings,H. and Motto,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-NOV-1997) Biologia Molecolare, ISC Bergamo, via
            Stezzano 24, Bergamo, I 24126, Italy
FEATURES             Location/Qualifiers
     source          1. .2047
                     /organism="Zea mays"
                     /strain="W22"
                     /db_xref="taxon:4577"
     gene            1. .2047
                     /gene="rpd3"
     CDS             141. .1682
                     /gene="rpd3"
                     /codon_start=1
                     /product="putative histone deacetylase RPD3"
                     /protein_id="AAC50038.1"
                     /db_xref="GI:2665840"
                     /translation="MDPSSAGSGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPM
                     KPHRIRMTHSLLARYGLLNQMQVYRPNPARERELCRFHAEEYINFLRSVTPETQQDQI
                     RLLKRFNVGEECPVLDGLYSFCQTYAGASVGGAVKFNHGHDIAINWSGGLHHAKKCEA
                     SGFCYVNDIVLAILELLKHHERVLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYF
                     PGTGDIRDIGHSKGKYYSLNVPLDDGIDDESYQSLFKPIMGKVMEVFRPGAVVLQCGA
                     DSLSGDRLGCFNLSIKGHAECVRYMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALG
                     QEPEDKMPVNEYYEYFGPDYTLHVAPSNMENKNTRQQLDDIRSKLSKLRHAPSVHFQE
                     RVPDTEIPEQDEDQDDPDERHDPDSDMEVDDHKAVEESSRRSILGIKIKREFGENATR
                     VQDGGRVASEHRGLEPMAEDIGSSKQAPQADASAMAIDEPSNVKNEPESSTKLQGQAA
                     AYHKP"
     polyA_site      1999
                     /gene="rpd3"
BASE COUNT      561 a    513 c    523 g    450 t
ORIGIN


  Query Match             39.3%;  Score 709.6;  DB 8;  Length 2047;
  Best Local Similarity   74.3%;  Pred. No. 1e-188;
  Matches  943;  Conservative    0;  Mismatches  309;  Indels   18;  Gaps    3;

Qy    144 ggatactggcggcaattcgctggcgtc---cggacctgatggtgtgaagaggaaagtttg 200
          ||   | ||||||||| || ||  ||||   ||| || || ||   |||  ||   || ||
Db    158 GGGCTCCGGCGGCAACTCCCTCCCGTCCGTCGGCCCCGACGGGCAGAAACGGCGCGTGTG 217

Qy    201 ttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatgaagcc 260
          || |||| |||| || || |||| |||||||| || || || |||| ||||||||
Db    218 CTACTTCTACGACCCGGATGTGGGCAACTACTACTACGGGCAGGGCCATCCGATGAAGCC 277

Qy    261 ccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcatatgca 320
          || |||||||| ||||| ||  | || || || | |||||| |||||  | || |||||
Db    278 GCACCGCATCCGGATGACGCACTCGCTGCTGGCGCGCTACGGCCTCCTCAACCAAATGCA 337

Qy    321 ggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgacta 380
          |||   |  |||  ||| ||||||||||| || ||||||||||||||||||||| || ||
Db    338 GGTGTACCGCCCCAACCCGGCCCGCGAACGCGAACTCTGCCGCTTCCACGCCGAGGAATA 397

Qy    381 tgtctcttttctccgcagcattacccctgaaaacccagcaagatcagattcgccaacttaa 440
            ||  || || |||  | | || || |||| |||| || |||| ||||  || ||
Db    398 CATCAACTTCCTGCGCTCCGTCACGCCGGAAACGCAGCAGGACCAGATCCGCCTGCTCAA 457

Qy    441 gcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagac 500
          |||||||| || || || || || |||||||| ||||||||||   ||| ||||||||
Db    458 GCGCTTCAACGTCGGCGAGGAGTGCCCCGTCTTGGACGGCCTTTACAGCTTCTGCCAGAC 517

Qy    501 ctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgc 560
          |||||| || |  || || || ||  | |||||| | ||||||||||     || || ||
Db    518 CTATGCGGGCGCCTCCGTCGGCGGGGCCGTCAAGTTCAACCACGGCC---ATGACATCGC 574

Qy    561 catcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgtta 620
          |||||||||  | || || || || |||| ||||||||||||||| || ||||||||||
Db    575 AATCAACTGGTCGGGGGGCCTGCACCACGCCAAGAAGTGCGAGGCGTCGGGCTTCTGTTA 634

Qy    621 cgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttcttta 680
          ||||||||| |||||  | || || || |||| || ||||| || ||| | |||| ||
Db    635 CGTCAATGACATCGTGCTCGCCATACTCGAGCTGCTCAAGCATCACGAGAGAGTTCTGTA 694

Qy    681 tgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctactgacag 740
          ||||||||| |||||||||| || || ||||||||||||||||| || ||  | || || ||
Db    695 TGTCGATATCGATATCCACCATGGTGATGGAGTGGAGGAGGCTTTCTACACAACAGATAG 754

Qy    741 ggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattca 800
          |||||||||||||||||| || || |||||||||| || || || |||||  | || ||
Db    755 GGTTATGACTGTCTCGTTCCACAAGTTTGGTGATTATTTCCCAGGAACAGGGGATATCCG 814

Qy    801 ggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaat 860
          || || ||  |   |  || |||||||| || || ||||| || || ||||||||| ||
Db    815 TGACATTGGGCACTCAAAAGGGAAGTACTACTCCCTGAATGTCCCTCTAGATGATGGGAT 874

Qy    861 cgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaaatttt 920
          |||||||| ||||| ||     | || ||||| |||||||| ||||||||||  ||||
Db    875 TGATGATGAAAGCTACCAGTCCCTTTTTAAGCCAATCATGGGCAAAGTTATGGAGGTTTT 934

Qy    921 ccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcggttagg 980
          ||| || || || || ||  | || |||||||||| ||| | ||||||||| |||| ||
Db    935 CCGCCCTGGTGCAGTTGTGCTTCAGTGTGGTGCTGATTCCTTGTCTGGGGATAGGTTGGG 994

Qy    981 ttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaa 1040
          |||||||| || ||||||||||||||||| || || || |  | |||||| || |||||
Db    995 CTGCTTCAACCTCTCAATCAAAGGTCATGCGGAATGTGTTAGGTATATGACGTCTTTCAA 1054
```

```
Qy   1041 tgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctg 1100
          |||||  |  ||||| | ||||||||||| || || ||  | |||||||| || || ||
Db   1055 CGTTCCATTGTTGCTCCTTGGTGGTGGTGGATATACCATAAGAAATGTTGCACGCTGTTG 1114

Qy   1101 gtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatga 1160
          ||| || ||||||||||||| |||||    ||   |||||||||||||| |   ||||
Db   1115 GTGTTATGAGACTGGAGTTGCTCTTGGCCAAGAGCCTGAAGACAAGATGCCTGTTAATGA 1174

Qy   1161 atattatgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaa 1220
          || |||||||||| |||||||| || || ||||| ||||| ||||||||||||| ||
Db   1175 GTACTATGAATACTTCGGTCCAGATTACACTCTTCATGTTGCACCAAGTAACATGGAGAA 1234

Qy   1221 taagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaa 1280
          || ||| | || ||    || || || || ||   |            || || || ||
Db   1235 CAAAAATACACGACAACAACTGGATGATATACGATCT----------AAACTTTCAAA 1282

Qy   1281 gcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagactcccga 1340
          ||||  |||||||| |||| |  ||||| || |||  ||||| ||||| | || ||
Db   1283 ACTTCGACATGCTCCTAGTGTCCACTTTCAAGAGAGAGTTCCTGACACAGAAATACCTGA 1342

Qy   1341 ggttgatgaagaccaagaagatggggataaaagatgggatccggattcagacatggatgt 1400
          |  |||||||| ||||| ||| ||| || ||  ||||| || || || ||||| ||
Db   1343 GCAAGATGAAGATCAAGATGATCCAGATGAACGACATGATCCTGACTCTGATATGGAAGT 1402

Qy   1401 tgatgatgac 1410
          |||||  ||
Db   1403 GGATGACCAC 1412
```

```
RESULT  11
AR168365
LOCUS       AR168365                2019 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 5 from patent US 6287843.
ACCESSION   AR168365
VERSION     AR168365.1  GI:17904243
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 2019)
  AUTHORS   Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
  TITLE     Maize histone deacetylases and their use
  JOURNAL   Patent: US 6287843-A 5 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .2019
                     /organism="unknown"
BASE COUNT      510 a    532 c    528 g    449 t
ORIGIN
```

```
  Query Match             38.7%;  Score 698.6;  DB 6;  Length 2019;
  Best Local Similarity   73.0%;  Pred. No. 1.3e-185;
  Matches  912;  Conservative    0;  Mismatches  334;  Indels    3;  Gaps    1;

Qy    162 gctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgaggt 221
          || | |  | ||  | || ||   |||| |    || || || ||||| ||| | |||||
Db    175 GCCGACCACTGGCGCGGACGGGTCGAAGCGCCGCGTCTGCTACTTCTACGACGCGGAGGT 234

Qy    222 cggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgaccca 281
          |||||  |||||||| || || || || || |||||||| || ||||||||||||||||
Db    235 GGGCAACTACTACTACGGGCAGGGCCACCCGATGAAGCCGCACCGCATCCGCATGACCCA 294

Qy    282 tgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagcccttccctgc 341
          || || ||||  | |||||| |||||  | || ||||| ||  ||   ||   ||||||
Db    295 CGCGCTGCTCGGCCGCTACGGCCTCCTCGACCAGATGCAAGTGTTCCGCCCTCACCCTGC 354

Qy    342 ccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagcat 401
          |||||| || || ||||||||||||||||||||||| || ||||| || |||||   | |
Db    355 CCGCGACCGCGACCTCTGCCGCTTCCACGCCGACGATTACGTCTCCTTCCTCCGGTCCGT 414

Qy    402 tacccctgaaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaaga 461
          ||||| ||||| ||||| || ||||| |||   || ||||||||||| || || || ||
```

```
Db    415 CACCCCCGAAACGCAGCAGGACCAGATCCGCGCGCTCAAGCGCTTCAACGTCGGCGAGGA 474

Qy    462 ctgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgttgg 521
          ||| |||||||| ||||| || ||    || || ||||| || || || || ||||||||
Db    475 CTGCCCCGTCTTCGACGGTCTCTACAGTTTCTGTCAGACGTACGCGGGGGGCTCTGTTGG 534

Qy    522 tggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggtct 581
          ||| | |||||||| ||||| ||||     ||||| |||||||||||||| || || ||
Db    535 CGGCGCCGTCAAGCTCAACCATGGCCAT---GATATCGCCATCAACTGGGCCGGCGGACT 591

Qy    582 ccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtcttagc 641
          ||| ||||| |||||||||||||||||| || || || || || ||||| || ||| | ||
Db    592 CCACCACGCCAAGAAGTGCGAGGCCTCCGGGTTTTGCTATGTTAATGACATTGTCCTCGC 651

Qy    642 tatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccacca 701
          ||||| |||||||| ||| | ||  |||| |||| || || || ||||||||||||||
Db    652 CATCCTCGAGCTCCTCAAGTACCACCAGCGCGTTCTGTACGTGGACATTGATATCCACCA 711

Qy    702 cggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgtttca 761
          |||||| || |||||||||||| |||||| | || ||| |||| ||||| ||||| || ||
Db    712 CGGGGACGGCGTGGAGGAGGCTTTTTATACCACAGACCGGGTGATGACAGTCTCATTCCA 771

Qy    762 taaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagcgg 821
          || ||||| ||||| || || || |||||  ||||||  ||| | ||  |    |  ||
Db    772 CAAGTTTGGAGATTATTTCCCTGGGACAGGGGACATTCGTGATGTTGGGCACTCAAAGGG 831

Qy    822 aaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcatct 881
          || || || || || ||||| || ||||| ||||| || ||||||||||||||| ||
Db    832 TAAATATTACTCCCTGAATGTTCCCCTGGACGATGGTATTGATGATGAGAGCTACCAGTC 891

Qy    882 gttattcaagcccatcatggggaaagttatggaaatttccgaccaggggctgtggtatt 941
          ||| |||||||| || ||||| || || |||||  | |||   || || || || ||  |
Db    892 GTTGTTCAAGCCAATAATGGGCAAGGTGATGGAGGTCTTCAACCCTGGTGCAGTCGTGCT 951

Qy    942 gcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatcaa 1001
          || |||||||| || ||  | || || ||  |||| || || |||| || || || ||
Db    952 CCAGTGTGGTGCGGATTCATTGTCGGGTGACAGGTTGGGCTGTTTCAACCTCTCTATTAA 1011

Qy   1002 aggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttggg 1061
          ||||| || || || || | |||||||| || ||||| || || ||  |||| | ||
Db   1012 GGGTCACGCAGAATGTGTAAGATTTATGAGGTCCTTCAACGTCCCGCTGTTGCTGCTTGG 1071

Qy   1062 tggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagttgc 1121
          ||||||||| || || ||  | || ||||| || || |||||||||||||| ||||||||
Db   1072 TGGTGGTGGGTATACCATAAGAAACGTTGCACGGTGTTGGTGCTACGAGACAGGAGTTGC 1131

Qy   1122 acttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggtcc 1181
          |||||  |||  |   ||||||||||     |||| || ||||| || ||||||||
Db   1132 CCTTGGTCATGAGCTCACTGACAAGATGCCACCTAATGAGTACTATGAGTATTTTGGTCC 1191

Qy   1182 agactatacacttcacgttgctccaagtaacatggaaaataagaatttctcgtcagatgct 1241
          ||| || || || || || |||||||||||||||| ||||| ||  | || ||    |
Db   1192 AGATTACACTCTACATGTCGCTCCAAGTAACATGGAGAATAAAAACACACGGCATCAATT 1251

Qy   1242 tgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagtgt 1301
          || || ||       | |||||  | ||||| || || || |  |||||||| |||||
Db   1252 GGATGACATAAAATCAAAACTTCTAGATAATCTTTCAAAACTCCGACATGCTCCTAGTGT 1311

Qy   1302 accatttcaggaaagaccacctgatacagagactcccgaggttgatgaagaccaagaaga 1361
          |  ||||| ||  |||| ||||| | |||   || |||   |||||||| ||||  |
Db   1312 TCAGTTTCAAGAGCGACCTCCTGAGGCTGAGTTACCTGAGCAAGATGAAGACAAAGAGAA 1371

Qy   1362 tggggataaaagatgggatccggattcagacatggatgttgatgatgac 1410
          |   ||| |||||   ||| | |||| ||  ||||  |  |||||| |
Db   1372 TCCTGATGAAAGACATGATGCTGATTCTGATGTGGAGATGAATGATGCC 1420
```

```
RESULT  12
AF045473
LOCUS       AF045473                1742 bp    mRNA    linear   PLN 28-JAN-1999
DEFINITION  Zea mays histone deacetylase (hd1b) mRNA, complete cds.
ACCESSION   AF045473
VERSION     AF045473.1  GI:4193319
```

```
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1742)
  AUTHORS   Pipal,A. and Wegener,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JAN-1998) Microbiology-Medical School, University of
            Innsbruck, Fritz-Pregl-Strasse 3, Innsbruck A-6020, Austria
FEATURES             Location/Qualifiers
     source          1. .1742
                     /organism="Zea mays"
                     /cultivar="Cuzco"
                     /db_xref="taxon:4577"
     gene            1. .1742
                     /gene="hdlb"
     CDS             1. .1377
                     /gene="hdlb"
                     /note="similar to Saccharomyces cerevisiae RPD3"
                     /codon_start=1
                     /product="histone deacetylase"
                     /protein_id="AAD10139.1"
                     /db_xref="GI:4193320"
                     /translation="MAASGEGVSLPSPAGGEDAHRRRVSYFYEPSIGDYYYGQGHPMK
                     PHRIRMAHSLVVHYGLHRLLELSRPYPASEADIRRFHSDDYVAFLASATGNPGVLDPR
                     AIKRFNVGEDCPVFDGLFPFCQASAGGSIGAAVKLNRGDADITVNWAGGLHHAKKSEA
                     SGFCYVNDIVLAILELLKFHRRVLYVDIDVHHGDGVEEAFFTTNRVMTVSFHKYGDFF
                     PGTGHITDVGAAEGKHYALNVPLSDGIDDTTFRGLFQCIIKKVMEVYQPDVVVLQCGA
                     DSLAGDRLGCFNLSVKGHADCLRFLRSYNVPMMVLGGGGYTIRNVARCWCYETAVAVG
                     VEPDNKLPYNDYYEYFGPDYTLHIQPKSVENLNTTKDLENIKNMILENLSKIEHVPST
                     QFHDRPSDPEAPEEKEEDMDKRPPQRSRLWSGGAYDSDTEDPDSLKSEGKDVTANFQM
                     KDEPKDDL"
     3'UTR           1378. .1742
                     /gene="hdlb"
BASE COUNT      407 a    457 c    453 g    425 t
ORIGIN
```

```
  Query Match           31.0%;  Score 559.4;  DB 8;  Length 1742;
  Best Local Similarity   66.8%;  Pred. No. 2.3e-146;
  Matches  797;  Conservative    0;  Mismatches  396;  Indels    0;  Gaps    0;

Qy    160 tcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgag 219
          || | ||||  |||    ||||          |    ||  | |||||||| || ||   |
Db     34 TCTCCGGCGGGCGGGGAGGATGCGCACCGCCGCCGCGTCAGCTATTTCTACGAGCCGTCG 93

Qy    220 gtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgacc 279
          ||||  | ||||||||| || |||||||| || ||||||||||| |||||||| ||| |
Db     94 ATCGGAGACTACTACTACGGGCAAGGTCACCCGATGAAGCCCCACCGCATCCGAATGGCG 153

Qy    280 catgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagcccttccct 339
          ||| | ||  | |  |||||||| ||||  |  |   |  || |   |   || | |||
Db    154 CATTCGCTGGTGGTCCACTACGGCCTCCACCGCCTCCTCGAGCTCTCCCGCCCGTACCCG 213

Qy    340 gcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagc 399
          |||   ||    ||  || |||||||||||| |||||||||| ||| |||| |||     |
Db    214 GCCTCTGAGGCCGACATCCGCCGCTTCCACTCCGACGACTACGTCGCTTTCCTCGCGTCC 273

Qy    400 attacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaa 459
             |||   | |           |   |||    |||||||||| || || |||||
Db    274 GCCACCGGAAACCCGGGTGTCCTCGACCCGCGCGCCATTAAGCGCTTTAACGTCGGTGAG 333

Qy    460 gactgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgtt 519
          |||||  |||||  || |||||  || |   ||||  ||||||  |||  |||||  |||    |
Db    334 GACTGCCCCGTGTTCGACGGTCTCTTCCCCTTCTGCCAGGCCTCCGCTGGGGGAAGCATC 393

Qy    520 ggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggt 579
          || | | | ||||||||||||| |||   |  ||| ||  ||  |||||||||| || ||
Db    394 GGCGCCGCCGTCAAGCTTAACCGCGGGGACGCCGACATCACCGTCAACTGGGCGGGCGGC 453

Qy    580 ctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtctta 639
```

```
          ||||| ||||| |||||| |||||||||| || ||||| |||||||| || |||||| |
Db    454 CTCCACCACGCCAAGAAGAGCGAGGCCTCCGGGTTCTGCTACGTCAACGACATCGTCCTC 513

Qy    640 gctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccac 699
          || ||||| |||||||| |||   ||   |||||| || ||||| || |||||| |||||
Db    514 GCCATCCTCGAGCTCCTCAAGTTCCACAGGCGTGTGCTATATGTAGACATTGATGTCCAC 573

Qy    700 cacggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgttt 759
          || || ||||| |||||||||||| || |   ||||  || | || |||||||| || |||
Db    574 CATGGAGATGGCGTGGAGGAGGCCTTCTTCACTACAAACCGAGTCATGACTGTTTCCTTT 633

Qy    760 cataaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagc 819
          || || | ||| ||||  || || ||||| || || ||    ||  | ||    |
Db    634 CACAAGTATGGGGATTTTTTCCCTGGTACTGGACATATCACTGACGTTGGGGCAGCCGAA 693

Qy    820 ggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcat 879
          || ||| | ||| |||| ||||| || ||| |||||| ||||||||    | || || |
Db    694 GGGAAGCATTATGCTCTGAATGTTCCCCTGAGTGATGGTATCGATGACACCACCTTTCGT 753

Qy    880 ctgttattcaagcccatcatggggaaagttatggaaatttccgaccaggggctgtggta 939
             | ||  |   ||||||   |||||||||||  |||  |  ||||  |  || ||
Db    754 GGTCTGTTTCAATGCATCATTAAGAAAGTTATGGAGGTTTATCAGCCAGACGTGGTTGTC 813

Qy    940 ttgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatc 999
           | ||||| || |||||||||  |  |||| ||  ||||||||||||||||| || ||  |
Db    814 CTCCAATGCGGAGCTGACTCTTTGGCTGGAGACAGGTTAGGTTGCTTCAACCTGTCTGTG 873

Qy   1000 aaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttg 1059
          || |||||||||| ||| ||   ||  | || |||| ||||||||  |  || | ||
Db    874 AAGGGTCATGCTGACTGCCTCCGTTTCCTTAGGTCGTACAATGTTCCTATGATGGTTTTA 933

Qy   1060 ggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagtt 1119
          |||||||| |||||||| ||| | |||||||| || ||||||||||||||||| | ||||
Db    934 GGTGGTGGAGGTTACACCATCAGAAATGTTGCACGCTGCTGGTGCTACGAGACCGCAGTT 993

Qy   1120 gcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggt 1179
          ||  |||||||||||  |||  ||||| ||||  |  |||| || ||||| ||||||||
Db    994 GCTGTTGGAGTTGAACCTGATAACAAGCTGCCTTACAATGATTACTATGAGTACTTTGGC 1053

Qy   1180 ccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatg 1239
          || || ||||| |||||  |   ||||  |   | |||||   ||||  |     ||
Db   1054 CCTGATTATACTCTTCATATCCAACCAAAAAGTGTTGAAAACCTGAATACCACAAAGGAC 1113

Qy   1240 cttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagt 1299
          | ||  | ||     ||     |  |  | || || || ||| |  | |||| ||| ||
Db   1114 TTGGAGAACATAAAGAACATGATATTGGAGAACCTGTCAAAGATAGAACATGTTCCCAGC 1173

Qy   1300 gtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaaga 1352
             | ||| || || |||||  | ||  | ||  |||| ||||   | || ||
Db   1174 ACTCAATTCCATGACAGACCGTCAGACCCTGAAGCTCCAGAGGAGAAAGAGGA 1226
```

```
RESULT  13
AR168363
LOCUS       AR168363                1826 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 1 from patent US 6287843.
ACCESSION   AR168363
VERSION     AR168363.1  GI:17904240
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1826)
  AUTHORS   Baldwin,D.Adelphi., Briggs,S.P. and Crane,V.C.
  TITLE     Maize histone deacetylases and their use
  JOURNAL   Patent: US 6287843-A 1 11-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .1826
                     /organism="unknown"
BASE COUNT      437 a    471 c    475 g    443 t
ORIGIN
```

```
  Query Match           31.0%;  Score 559.4;  DB 6;  Length 1826;
  Best Local Similarity  66.8%;  Pred. No. 2.4e-146;
  Matches  797;  Conservative    0;  Mismatches  396;  Indels    0;  Gaps    0;

Qy    160 tcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatgaccctgag 219
          || | |||| ||| |||| |          |  ||  | |||||||| || ||   |
Db     62 TCTCCGGCGGGCGGGGAGGATGCGCACCGCCGCCGCGTCAGCTATTTCTACGAGCCGTCG 121

Qy    220 gtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatccgcatgacc 279
          ||||  | |||||||| || |||||||| || ||||||||||| |||||||| ||| |
Db    122 ATCGGAGACTACTACTACGGGCAAGGTCACCCGATGAAGCCCCACCGCATCCGAATGGCG 181

Qy    280 catgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagcccttccct 339
          ||  | ||  | |  |||||||| |||| |  |   |  || |   |  |||| |||
Db    182 CACTCGCTGGTGGTCCACTACGGCCTCCACCGCCTCCTCGAGCTCTCCCGCCCCTACCCG 241

Qy    340 gcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttctccgcagc 399
          |||   ||    ||  || |||||||||||| |||||||||| ||| |||| |||    |
Db    242 GCCTCTGAGGCCGACATCCGCCGCTTCCACTCCGACGACTACGTCGCTTTCCTCGCGTCC 301

Qy    400 attacccctgaaacccagcaagatcagattcgccaacttaagcgcttcaatgttggtgaa 459
          |||   |  |           |   |||    |||||||||| || || |||||
Db    302 GCCACCGGAAACCCGGGTGTCCTCGACCCGCGCGCCATTAAGCGCTTTAACGTCGGTGAG 361

Qy    460 gactgtcccgtctttgacggcctttattccttttgccagacctatgctggaggatctgtt 519
          ||||| ||||| || ||||| || |   |||| |||||| |||  ||||| |||    |
Db    362 GACTGCCCCGTGTTCGACGGTCTCTTCCCCTTCTGCCAGGCCTCCGCTGGGGGAAGCATC 421

Qy    520 ggtggctctgtcaagcttaaccacggcctctgcgatattgccatcaactgggctggtggt 579
          || | | |||||||||||||| |||   |  ||| ||  || ||||||||||| || ||
Db    422 GGCGCCGCCGTCAAGCTTAACCGCGGGGACGCCGACATCACCGTCAACTGGGCGGGCGGC 481

Qy    580 ctccatcacgctaagaagtgcgaggcctctggcttctgttacgtcaatgatatcgtctta 639
          ||||| ||||| |||||| |||||||||| || ||||| |||||||| || ||||||  |
Db    482 CTCCACCACGCCAAGAAGAGCGAGGCCTCCGGGTTCTGCTACGTCAACGACATCGTCCTC 541

Qy    640 gctatcctagagctccttaagcagcatgagcgtgttctttatgtcgatattgatatccac 699
          || ||||| |||||||| |||   ||   |||||| || ||||| || |||||| |||||
Db    542 GCCATCCTCGAGCTCCTCAAGTTCCACAGGCGTGTGCTATATGTAGACATTGATGTCCAC 601

Qy    700 cacggggatggagtggaggaggcattttatgctactgacagggttatgactgtctcgttt 759
          || || ||||| |||||||||||| || |   ||||  || | || |||||||| || |||
Db    602 CATGGAGATGGCGTGGAGGAGGCCTTCTTCACTACAAACCGAGTCATGACTGTTTCCTTT 661

Qy    760 cataaatttggtgattactttcccggtacaggtcacattcaggatataggttatggtagc 819
          || || | ||| |||| || || ||||| || || ||    ||  | ||    |
Db    662 CACAAGTATGGGGATTTTTTCCCTGGTACTGGACATATCACTGACGTTGGGGCAGCCGAA 721

Qy    820 ggaaagtactattctctcaatgtaccactggatgatggaatcgatgatgagagctatcat 879
          || ||| | ||| |||| ||||| || |||  |||||| ||||||||    | || || |
Db    722 GGGAAGCATTATGCTCTGAATGTTCCCCTGAGTGATGGTATCGATGACACCACCTTTCGT 781

Qy    880 ctgttattcaagcccatcatggggaaagttatggaaattttccgaccaggggctgtggta 939
          | ||  |  ||||||  |||||||||||| |||  |  |||| |  || ||
Db    782 GGTCTGTTTCAATGCATCATTAAGAAAGTTATGGAGGTTTATCAGCCAGACGTGGTTGTC 841

Qy    940 ttgcaatgtggtgctgactccctatctggggatcggttaggttgcttcaatctttcaatc 999
          | ||||| || |||||||  |  |||| ||  ||||||||||||||| || ||  |
Db    842 CTCCAATGCGGAGCTGACTCTTTGGCTGGAGACAGGTTAGGTTGCTTCAACCTGTCTGTG 901

Qy   1000 aaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttcccctactgctcttg 1059
          || |||||||||| ||| ||   ||  | || |||| ||||||||  |  || | ||
Db    902 AAGGGTCATGCTGACTGCCTCCGTTTCCTTAGGTCGTACAATGTTCCTATGATGGTTTTA 961

Qy   1060 ggtggtggtggttacactatccgcaatgttgcccgttgctggtgctacgagactggagtt 1119
          |||||||| |||||||| ||| | |||||||| || ||||||||||||||||  | ||||
Db    962 GGTGGTGGAGGTTACACCATCAGAAATGTTGCACGCTGCTGGTGCTACGAGACCGCAGTT 1021

Qy   1120 gcacttggagttgaagttgaagacaagatgccggagcatgaatattatgaatactttggt 1179
          ||  |||||||||| ||| ||||| |||   |  |||| || ||||| ||||||||
Db   1022 GCTGTTGGAGTTGAACCTGATAACAAGCTGCCTTACAATGATTACTATGAGTACTTTGGC 1081

Qy   1180 ccagactatacacttcacgttgctccaagtaacatggaaaataagaattctcgtcagatg 1239
```

```
          || || |||| |||||  |    ||||  |   | |||||   |||| |     ||
Db   1082 CCTGATTATACTCTTCATATCCAACCAAAAAGTGTTGAAAACCTGAATACCACAAAGGAC 1141

Qy   1240 cttgaagagattcgcaatgaccttctccacaatctctctaagcttcagcatgctccaagt 1299
          | ||  | ||    ||     |  |  | || || || ||| |  | |||| ||| ||
Db   1142 TTGGAGAACATAAAGAACATGATATTGGAGAACCTGTCAAAGATAGAACATGTTCCCAGC 1201

Qy   1300 gtaccatttcaggaaagaccacctgatacagagactcccgaggttgatgaaqa 1352
             | ||| || || |||||  | ||  | ||  |||| ||||   | || ||
Db   1202 ACTCAATTCCATGACAGACCGTCAGACCCTGAAGCTCCAGAGGAGAAAGAGGA 1254


RESULT  14
AF440226
LOCUS       AF440226               1662 bp    mRNA    linear   PLN 20-NOV-2001
DEFINITION  Zea mays histone deacetylase (HDA108) mRNA, partial cds.
ACCESSION   AF440226
VERSION     AF440226.1  GI:17017397
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 1662)
  AUTHORS   Chandler,V.L., Kaeppler,S.M., Kaeppler,H.F. and Cone,K.C.
  TITLE     Sequences from the Plant Chromatin Consortium (NSF Plant Genome
            project 9975930)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1662)
  AUTHORS   Bergstrom,D., Springer,N.M., Schmitt,L.T., Guthrie,E.,
            Sidorenko,L., Selinger,D., Kaeppler,S.M. and Cone,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-OCT-2001) Biological Sciences, University of
            Missouri, 101 Tucker Hall, Columbia, MO 65211, USA
FEATURES             Location/Qualifiers
     source          1. .1662
                     /organism="Zea mays"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
     gene            1. .1662
                     /gene="HDA108"
     CDS             <1. .1311
                     /gene="HDA108"
                     /note="similar to RPD3, HD2 and Sir2"
                     /codon_start=1
                     /product="histone deacetylase"
                     /protein_id="AAL33653.1"
                     /db_xref="GI:17017398"
                     /translation="RVSYFYEPSIGDYYYGQGHPMKPHRIRMAHSLVVHYGLHRLLEL
                     SRPYPASEADIRRFHSDDYVAFLASATGNPGVLDPRAIKRFNVGEDCPVFDGLFPFCQ
                     ASAGGSIGAAVKLNRGDADITVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKFHRR
                     VLYVDIDVHHGDGVEEAFFTTNRVMTVSFHKYGDFFPGTGHITDVGAAEGKHYALNVP
                     LSDGIDDTTFRGLFQCIIKKVMEVYQPDVVVLQCGADSLAGDRLGCFNLSVKGHADCL
                     RFLRSYNVPMMVLGGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNDYYEYFGPDYTL
                     HIQPKSVENLNTTKDLENIKNMILENLSKIEHVPSTQFHDRPSDPEAPEEKEEDMDKR
                     PPQRSRLWSGGAYDSDTEDPDSLKSEGKDVTANFQMKDEPKDDL"
BASE COUNT      396 a    438 c    421 g    407 t
ORIGIN


  Query Match           30.8%;  Score 557;  DB 8;  Length 1662;
  Best Local Similarity  67.6%;  Pred. No. 1.1e-145;
  Matches  782;  Conservative    0;  Mismatches  375;  Indels    0;  Gaps    0;

Qy    196 gtttgttatttctatgaccctgaggtcggcaattactactatggccaaggtcatcccatg 255
          ||  | |||||||| || ||   | ||||  | |||||||| || |||||||| || |||
Db      4 GTCAGCTATTTCTACGAGCCGTCGATCGGAGACTACTACTACGGGCAAGGTCACCCGATG 63

Qy    256 aagccccatcgcatccgcatgacccatgccctcctcgctcactacggtctccttcagcat 315
          |||||||| |||||||| ||| | ||  | ||  | |  |||||||| ||||  |  |
Db     64 AAGCCCCACCGCATCCGAATGGCGCACTCGCTGGTGGTCCACTACGGCCTCCACCGCCTC 123
```

```
Qy    316 atgcaggttctcaagcccttccctgcccgcgaacgtgatctctgccgcttccacgccgac 375
          |  || |   |   |||| ||| |||   ||    ||  || ||||||||||| |||||
Db    124 CTCGAGCTCTCCCGCCCCTACCCGGCCTCTGAGGCCGACATCCGCCGCTTCCACTCCGAC 183

Qy    376 gactatgtctcttttctccgcagcattacccctgaaacccagcaagatcagattcgccaa 435
          ||||| ||| |||| |||     |   |||     |  |           |    |||
Db    184 GACTACGTCGCTTTCCTCGCGTCCGCCACCGGAAACCCGGGTGTCCTCGACCCGCGCGCC 243

Qy    436 cttaagcgcttcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgc 495
          ||||||||||| || || ||||| ||||| ||||| || ||||| || |   |||| |||
Db    244 ATTAAGCGCTTTAACGTCGGTGAGGACTGCCCCGTGTTCGACGGTCTCTTCCCCTTCTGC 303

Qy    496 cagacctatgctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgat 555
          ||| |||  ||||| |||     | || | | ||||||||||||| |||   |  |||
Db    304 CAGGCCTCCGCTGGGGGAAGCATCGGCGCCGCCGTCAAGCTTAACCGCGGGGACGCCGAC 363

Qy    556 attgccatcaactgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttc 615
          ||  || |||||||||| || || ||||| ||||| |||||| |||||||||| || |||
Db    364 ATCACCGTCAACTGGGCGGGCGGCCTCCACCACGCCAAGAAGAGCGAGGCCTCCGGGTTC 423

Qy    616 tgttacgtcaatgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgtt 675
          || |||||||| || |||||| | || ||||| |||||||| |||   ||   ||||||
Db    424 TGCTACGTCAACGACATCGTCCTCGCCATCCTCGAGCTCCTCAAGTTCCACAGGCGTGTG 483

Qy    676 ctttatgtcgatattgatatccaccacggggatggagtggaggaggcattttatgctact 735
          || ||||| || |||||| |||||| || ||||| |||||||||| || |   ||||
Db    484 CTATATGTAGACATTGATGTCCACCATGGAGATGGCGTGGAGGAGGCCTTCTTCACTACA 543

Qy    736 gacagggttatgactgtctcgtttcataaatttggtgattactttcccggtacaggtcac 795
          || | || |||||||| || ||||| || | ||| |||  || || ||||| || ||
Db    544 AACCGAGTCATGACTGTTTCCTTTCACAAGTATGGGGATTTTTTCCCTGGTACTGGACAT 603

Qy    796 attcaggatataggttatggtagcggaaagtactattctctcaatgtaccactggatgat 855
          ||    ||  | ||    |     || ||| | ||| |||| ||||| || |||  ||||
Db    604 ATCACTGACGTTGGGGCAGCCGAAGGGAAGCATTATGCTCTGAATGTTCCCCTGAGTGAT 663

Qy    856 ggaatcgatgatgagagctatcatctgttattcaagcccatcatggggaaagttatggaa 915
          || ||||||||    | || || |    | || |   ||||||   ||||||||||||
Db    664 GGTATCGATGACACCACCTTTCGTGGTCTGTTTCAATGCATCATTAAGAAAGTTATGGAG 723

Qy    916 attttccgaccaggggctgtggtattgcaatgtggtgctgactccctatctggggatcgg 975
          |||  |  |||| |  || ||  | ||||| || |||||||  |  |||| ||  ||
Db    724 GTTTATCAGCCAGACGTGGTTGTCCTCCAATGCGGAGCTGACTCTTTGGCTGGAGACAGG 783

Qy    976 ttaggttgcttcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcg 1035
          |||||||||||||| || ||  | || |||||||||| ||| ||   ||  | || |||
Db    784 TTAGGTTGCTTCAACCTGTCTGTGAAGGGTCATGCTGACTGCCTCCGTTTCCTTAGGTCG 843

Qy   1036 ttcaatgttcccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgt 1095
          | |||||||||  |  || | || |||||||| |||||||| ||| | |||||||| ||
Db    844 TACAATGTTCCTATGATGGTTTTAGGTGGTGGAGGTTACACCATCAGAAATGTTGCACGC 903

Qy   1096 tgctggtgctacgagactggagttgcacttggagttgaagttgaagacaagatgccggag 1155
          |||||||||||||||| | ||||||  |||||||||  |||  ||||| |||| |
Db    904 TGCTGGTGCTACGAGACCGCAGTTGCTGTTGGAGTTGAACCTGATAACAAGCTGCCTTAC 963

Qy   1156 catgaatattatgaatactttggtccagactatacacttcacgttgctccaagtaacatg 1215
          |||| || ||||| |||||||| || || ||||| ||||| |    ||||  |   |
Db    964 AATGATTACTATGAGTACTTTGGCCCTGATTATACTCTTCATATCCAACCAAAAAGTGTT 1023

Qy   1216 gaaaataagaattctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctc 1275
          |||||  |||| |     ||    | ||  | ||    ||      |  |  | || ||
Db   1024 GAAAACCTGAATACCACAAAGGACTTGGAGAACATAAAGAACATGATATTGGAGAACCTG 1083

Qy   1276 tctaagcttcagcatgctccaagtgtaccatttcaggaaagaccacctgatacagagact 1335
          || ||| |  | |||| ||| ||    | ||| || || |||||  | ||  | ||  ||
Db   1084 TCAAAGATAGAACATGTTCCCAGCACTCAATTCCATGACAGACCGTCAGACCCTGAAGCT 1143

Qy   1336 cccgaggttgatgaaga 1352
          || ||||   | || ||
Db   1144 CCAGAGGAGAAAGAGGA 1160
```

```
RESULT  15
AY072201
LOCUS       AY072201                1704 bp    mRNA    linear   PLN 16-JAN-2002
DEFINITION  Arabidopsis thaliana putative histone deacetylase (At5g63110) mRNA,
            complete cds.
ACCESSION   AY072201
VERSION     AY072201.1  GI:18176318
KEYWORDS    FLI_CDNA.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1704)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Full Length cDNA of gene At5g63110 (GI:15242626)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1704)
  AUTHORS   Yamada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,
            Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,
            Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,
            Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,
            Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
            Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
            Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,
            Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2002) Plant Gene Expression Center, 800 Buchanan
            Street, Albany, CA 94710, USA
COMMENT     RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Yamada,K., Banh,J.,
            Chan,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,
            Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,
            Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,
            Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,
            Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,
            Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.

            Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to
            this work.  Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)
            contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1704
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="RAFL09-59-F23 (R12957)"
                     /note="This clone is in a modified pBluescript vector
                     (FLC-1) as a BamHI/XhoI insert.
                     ecotype: Columbia"
     gene            1. .1704
                     /gene="At5g63110"
     5'UTR           1. .75
                     /gene="At5g63110"
     CDS             76. .1491
                     /gene="At5g63110"
                     /codon_start=1
                     /evidence=experimental
                     /product="putative histone deacetylase"
                     /protein_id="AAL60022.1"
```

```
                     /db_xref="GI:18176319"
                     /translation="MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHR
                     IRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARN
                     LRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEAS
                     GFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFP
                     GTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGAD
                     SLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGGGYTIRNVARCWCYETAVAVGV
                     EPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQ
                     FQHTPPVNRVLDEPEDDMETRPKPRIWSGTATYESDSDDDDKPLHGYSCRGGATTDRD
                     STGEDEMDDDNPEPDVNPPSS"
    3'UTR            1492. .1704
                     /gene="At5g63110"
BASE COUNT      454 a    363 c    433 g    454 t
ORIGIN


  Query Match           29.1%;  Score 525;  DB 8;  Length 1704;
  Best Local Similarity  66.0%;  Pred. No. 1.2e-136;
  Matches  778;  Conservative    0;  Mismatches  395;  Indels    6;  Gaps    1;

Qy    152 gcggcaattcgctggcgtccggacctgatggtgtgaagaggaaagtttgttatttctatg 211
          ||||||  || ||| |||| ||||| || ||    ||| ||  |||  |||| ||||| |
Db     92 GCGGCATCTCTCTGCCGTCGGGACCCGACGGACGTAAGCGGCGAGTCAGTTACTTCTACG 151

Qy    212 accctgaggtcggcaattactactatggccaaggtcatcccatgaagccccatcgcatcc 271
          | ||   | ||||  | |||||||| || ||||| || || |||||||| || || ||||
Db    152 AGCCGACGATCGGAGACTACTACTACGGTCAAGGCCACCCGATGAAGCCTCACCGGATCC 211

Qy    272 gcatgacccatgccctcctcgctcactacggtctccttcagcatatgcaggttctcaagc 331
          | ||| | |||  ||| ||  ||||||    |||| |  |   |  |  |       |
Db    212 GTATGGCTCATAGCCTAATCATTCACTATCACCTCCACCGTCGCTTAGAAATCAGTCGCC 271

Qy    332 ccttccctgcccgcgaacgtgatctctgccgcttccacgccgacgactatgtctcttttc 391
          |   ||  ||   ||     ||| || |||| |||||  |    || |||||    || |
Db    272 CTAGCCTCGCTGACGCCTCCGATATCGGCCGATTCCATTCGCCGGAGTATGTTGACTTCC 331

Qy    392 tccgcagcattacccctgaaacccagcaagatc------agattcgccaacttaagcgct 445
          ||     | || | || ||| |   |   ||||           ||  | || | ||| |
Db    332 TCGCTTCCGTTTCGCCGGAATCTATGGGCGATCCTTCCGCTGCACGAAACCTAAGGCGAT 391

Qy    446 tcaatgttggtgaagactgtcccgtctttgacggcctttattccttttgccagacctatg 505
          ||||||| ||||| || ||||| ||||| ||||| |||| |   |||||||   | |  |
Db    392 TCAATGTCGGTGAGGATTGTCCTGTCTTCGACGGACTTTTTGATTTTTGCCGTGCTTCCG 451

Qy    506 ctggaggatctgttggtggctctgtcaagcttaaccacggcctctgcgatattgccatca 565
          | ||||| ||| ||||||   | |||||  | |||         |   ||||| || ||||
Db    452 CCGGAGGTTCTATTGGTGCTGCCGTCAAATTAAACAGACAGGACGCTGATATCGCTATCA 511

Qy    566 actgggctggtggtctccatcacgctaagaagtgcgaggcctctggcttctgttacgtca 625
          | ||||  ||||| || || || ||||||||  ||||||| ||||| || || || || |
Db    512 ATTGGGGCGGTGGGCTTCACCATGCTAAGAAAAGCGAGGCTTCTGGGTTTTGCTATGTAA 571

Qy    626 atgatatcgtcttagctatcctagagctccttaagcagcatgagcgtgttctttatgtcg 685
          | || |||||  |||  || || ||| | || |||  |  | |||| ||||| ||  | |
Db    572 ACGACATCGTGCTAGGGATTCTGGAGTTGCTCAAGATGTTTAAGCGGGTTCTCTACATAG 631

Qy    686 atattgatatccaccacggggatggagtggaggaggcattttatgctactgacagggtta 745
          |||||||| ||||||| || |||||||||| || || |||||  | ||||| || ||||
Db    632 ATATTGATGTCCACCATGGAGATGGAGTGGAAGAAGCGTTTTACACCACTGATAGAGTTA 691

Qy    746 tgactgtctcgtttcataaatttggtgattactttcccggtacaggtcacattcaggata 805
          |||||||  || || || |||||||| || |  || || || || ||||||||    |||
Db    692 TGACTGTTTCTTTCCACAAATTTGGGGACTTTTTCCCAGGAACTGGTCACATAAGAGATG 751

Qy    806 taggttatggtagcggaaagtactattctctcaatgtaccactggatgatggaatcgatg 865
          | ||   ||  |  || || |||||| |||| ||||| ||||  | ||||| || || |
Db    752 TTGGCGCTGAAAAAGGGAAATACTATGCTCTAAATGTTCCACTAAACGATGGTATGGACG 811

Qy    866 atgagagctatcatctgttattcaagcccatcatggggaaagttatggaaattttccgac 925
          |||| || | |     || || |  ||  | ||   ||| |||||||| | |  |  |
Db    812 ATGAAAGTTTCCGCAGCTTGTTTAGACCTCTTATCCAGAAGGTTATGGAAGTGTATCAGC 871

Qy    926 cagggctgtggtattgcaatgtggtgctgactccctatctggggatcggttaggttgct 985
          '
```

```
          ||| ||| || ||  | || ||||||||||||||| ||  ||| |||||||| |||||||
Db    872 CAGAGGCAGTTGTTCTTCAGTGTGGTGCTGACTCCTTAAGTGGTGATCGGTTGGGTTGCT 931

Qy    986 tcaatctttcaatcaaaggtcatgctgagtgcgtcaaatttatgagatcgttcaatgttc 1045
          ||||  | ||| |||| ||||| ||||| ||| |    ||  | ||||| | ||| ||||
Db    932 TCAACTTATCAGTCAAGGGTCACGCTGATTGCCTTCGGTTCTTAAGATCTTACAACGTTC 991

Qy   1046 ccctactgctcttgggtggtggtggttacactatccgcaatgttgcccgttgctggtgct 1105
          | ||  || | |||||||||||| || || ||||| || |||||||||||||||||||| |
Db    992 CTCTCATGGTGTTGGGTGGTGGAGGGTATACTATTCGAAATGTTGCCCGTTGCTGGTGTT 1051

Qy   1106 acgagactggagttgcacttggagttgaagttgaagacaagatgccggagcatgaatatt 1165
          | ||||||| |||||  ||||||| ||    ||  ||||  | ||  |  |||| ||||
Db   1052 ATGAGACTGCAGTTGCTGTTGGAGTAGAGCCGGACAACAAACTCCCTTACAATGAGTATT 1111

Qy   1166 atgaatactttggtccagactatacacttcacgttgctccaagtaacatggaaaataaga 1225
           ||| || || || ||||| ||||| ||||| || |  ||||||   ||||| |||   |
Db   1112 TTGAGTATTTCGGCCCAGATTATACGCTTCATGTCGACCCAAGTCCTATGGAGAATTTAA 1171

Qy   1226 attctcgtcagatgcttgaagagattcgcaatgaccttctccacaatctctctaagcttc 1285
          |  | |   |     | ||   |||  | ||     | ||  |  | || ||    ||
Db   1172 ACACGCCCAAAGATATGGAGAGGATAAGGAACACGTTGCTGGAACAACTTTCGGGACTAA 1231

Qy   1286 agcatgctccaagtgtaccatttcaggaaagaccacctg 1324
            || || || || || |  |||||| | | |||||| |
Db   1232 TACACGCACCTAGCGTCCAGTTTCAGCACACACCACCAG 1270


Search completed: April 28, 2002, 19:26:38
Job time: 9248 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:51:32 ; Search time 44.51 Seconds
(without alignments)
1646.426 Million cell updates/sec

Title: US-09-645-337-2
Perfect score: 2722
Sequence: 1 MDTGGNSLASGPDGVKRKVC.....KMEEEGTNKGGAEQAFPPKT 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2719	99.9	501	10	Q9SZL3	Q9szl3 arabidopsis
2	2134.5	78.4	500	10	Q9LKG1	Q9lkg1 mesembryant
3	1958.5	72.0	493	10	Q9AXF0	Q9axf0 oryza sativ
4	1568.5	57.6	471	10	Q9FML2	Q9fml2 arabidopsis
5	1561.5	57.4	471	10	Q9FVE5	Q9fve5 arabidopsis
6	1548	56.9	458	10	Q9ZTP8	Q9ztp8 zea mays (m
7	1496.5	55.0	521	5	O77213	O77213 drosophila
8	1496.5	55.0	521	5	Q9VZA1	Q9vza1 drosophila
9	1374	50.5	428	11	Q99PA0	Q99pa0 rattus norv
10	1361.5	50.0	687	3	Q9P4F5	Q9p4f5 emericella
11	1359	49.9	428	11	Q9JM08	Q9jm08 mus musculu
12	1358	49.9	428	11	Q9JLX5	Q9jlx5 mus musculu
13	1336	49.1	405	3	O59702	O59702 schizosacch
14	1334	49.0	465	5	O62339	O62339 caenorhabdi
15	1329.5	48.8	648	3	Q9C1C6	Q9c1c6 cochliobolu
16	1320.5	48.5	444	5	Q9GUA8	Q9gua8 cryptospori
17	1313	48.2	438	5	Q9VNC2	Q9vnc2 drosophila
18	1307	48.0	449	5	Q9XYC7	Q9xyc7 plasmodium
19	1298	47.7	419	10	Q9M1N6	Q9m1n6 arabidopsis
20	1259	46.3	566	3	Q9HDT2	Q9hdt2 ustilago ma
21	1240	45.6	437	5	Q9GU59	Q9gu59 cryptospori
22	1194	43.9	409	10	Q9FH09	Q9fh09 arabidopsis
23	1138	41.8	429	5	Q9GT74	Q9gt74 tetrahymena
24	1006	37.0	481	3	Q9P4F4	Q9p4f4 emericella
25	852	31.3	377	4	Q9NP76	Q9np76 homo sapien
26	847	31.1	377	4	Q9NYH4	Q9nyh4 homo sapien
27	847	31.1	377	11	Q9D0K6	Q9d0k6 mus musculu
28	759.5	27.9	428	5	Q9GRP1	Q9grp1 leishmania
29	644.5	23.7	256	11	Q99PA2	Q99pa2 rattus norv
30	594.5	21.8	223	4	Q9H368	Q9h368 homo sapien
31	475	17.5	375	2	O67135	O67135 aquifex aeo
32	448.5	16.5	389	2	Q9K7X1	Q9k7x1 bacillus ha
33	425.5	15.6	367	2	Q9WX04	Q9wx04 streptomyce
34	389.5	14.3	389	2	Q99TC9	Q99tc9 staphylococ
35	367.5	13.5	158	10	Q9LXN8	Q9lxn8 arabidopsis
36	340	12.5	883	5	Q9XYX1	Q9xyx1 drosophila
37	317.5	11.7	380	2	Q9HXM1	Q9hxm1 pseudomonas
38	309	11.4	1108	10	Q9FNQ7	Q9fnq7 arabidopsis
39	301.5	11.1	425	10	Q9SJE6	Q9sje6 arabidopsis
40	299.5	11.0	142	10	Q9M1N8	Q9m1n8 arabidopsis
41	298.5	11.0	1215	4	O94975	O94975 homo sapien
42	297	10.9	310	2	O67877	O67877 aquifex aeo
43	296.5	10.9	1063	4	Q9NSW6	Q9nsw6 homo sapien
44	296.5	10.9	1066	4	Q9NZS3	Q9nzs3 homo sapien
45	293	10.8	577	10	Q9LS38	Q9ls38 arabidopsis

ALIGNMENTS

RESULT 1
Q9SZL3
ID Q9SZL3 PRELIMINARY; PRT; 501 AA.
AC Q9SZL3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HISTONE DEACETYLASE.
GN F20D10.250 OR AT4G38130.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Wedler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Wedler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL035538; CAB37553.1; -.
DR EMBL; AL161593; CAB80478.1; -.
DR InterPro; IPR000286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 501 AA; 56023 MW; 857D2E3D16B7CC1F CRC64;

Query Match 99.9%; Score 2719; DB 10; Length 501;
Best Local Similarity 99.8%; Pred. No. 5.3e-205;

```
  Matches  500;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
           ||||||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VLKPFPARDRDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120

Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360

Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420

Qy     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480

Qy     481 VKMEEEGTNKGGAEQAFPPKT 501
           |||||||||||||||||||||
Db     481 VKMEEEGTNKGGAEQAFPPKT 501


RESULT   2
Q9LKG1
ID   Q9LKG1      PRELIMINARY;      PRT;   500 AA.
AC   Q9LKG1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
GN   HDEAC1.
OS   Mesembryanthemum crystallinum (Common ice plant).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX   NCBI_TaxID=3544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Scharte J., Baur B.;
RT   "Molecular cloning of histone deacetylase from Mesembryanthemum
RT   crystallinum.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF282858; AAF82385.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   500 AA;  56264 MW;  CFD4FE0525209ABC CRC64;


  Query Match             78.4%;  Score 2134.5;  DB 10;  Length 500;
  Best Local Similarity   79.8%;  Pred. No. 3.5e-159;
  Matches  399;  Conservative   31;  Mismatches   61;  Indels    9;  Gaps    3;

Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
           |||||||| |||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1 MDTGGNSLPSGPDGVKRKVSYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
           |||| |||:|||||||||||||||| |||| ||||:|||||||||||||||||||||||||
Db      61 VLKPVPARDRDLCRFHADDYVSFLRGITPEMQQDQLRQLKRFNVGEDCPVFDGLYSFCQT 120

Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           ||||||||:|||||||||||||||||||||| |||||||||||||||||| ||||||
Db     121 YAGGSVGGAVKLNHGLCDIAINWAGGLHHAKKSEASGFCYVNDIVLAILELLKVHERVLY 180

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
           |||||||||||||||| ||:|||||||||||||||||      |||| ||||||||||||||
Db     181 VDIDIHHGDGVEEAFYTTERVMTVSFHKFGDYFPGTETFVYIGYGKGKYYSLNVPLDDGI 240

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
           |||||  |||||||||||||:|||||||||||||||||||||||||||||||||||:| |||
Db     241 DDESYQSLFKPIMGKVMEIFKPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKYMISFN 300

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
           |||||:|||||||| |||||||||||||| ||::||:|:|||||||||||||||||||||||
Db     301 VPLLLVGGGGYTIRYVARCWCYETGVALGFELDDKIPQHEYYEYFGPDYTLHVAPSNMEN 360

Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           |||| ||::|| || || ||||||| |||||: | || |||| | : ||||||||:|
Db     361 KNSRPMLDDIRGKLLEYLSSLQHAPSVQFQERPPEAEIPEEDEDQHDPEGRWDPDSDMEV 420

Qy     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVE-EA 479
           ||    | |||||| :||| ||:| :|  ::  :  :|    :|:| |        : |
Db     421 DDRSISIRSRVKREIMEPDLKDRDDVKDAVDHPRDMDVSYG-AGNTSVKAPTTSSDQPEG 479

Qy     480 SVKMEEEGTNKGGAEQAFPP 499
             ||:: :|:      ||
Db     480 CVEMEQDNSNR-------PP 492


RESULT   3
Q9AXF0
ID   Q9AXF0      PRELIMINARY;      PRT;   493 AA.
AC   Q9AXF0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE HD1.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. YUANFENGZAO;
RA   Song F., Goodman R.M.;
RT   "Molecular characterization of a rice histone deacetylase gene
RT   OsHD1.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF332875; AAK01712.1; -.
SQ   SEQUENCE   493 AA;  55054 MW;  0EE300BB42243C13 CRC64;


  Query Match             72.0%;  Score 1958.5;  DB 10;  Length 493;
  Best Local Similarity   77.5%;  Pred. No. 2.2e-145;
  Matches  371;  Conservative   36;  Mismatches   59;  Indels   13;  Gaps    5;

Qy       4 GGNSL-ASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVL 62
           |||||  :| || ||:||:||| |||||| ||||||||||||||||||||||||  ||||
Db       8 GGNSLPTAGADGAKRRVCHFYDAEVGNYYCGQGHPMKPHRIRMTHALLAHYGLLDQMQVL 67

Qy      63 KPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYA 122
           || |||:|||||||||||:||||:|||||||||| ||||||||||||||||||||||||||
Db      68 KPHPARDRDLCRFHADDYVAFLRSVTPETQQDQIRALKRFNVGEDCPVFDGLYSFCQTYA 127

Qy     123 GGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVD 182
           ||||||:|||||||  |||||||||||||||||||||||||||||||||||||||| |:|||||
```

```
Db    128 GGSVGGAVKLNHG-HDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKYHQRVLYVD 186

Qy    183 IDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDD 242
          |||||||||||||| ||||||||||||||||||| |:|||| |||| ||||||||||
Db    187 IDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGYSEGKYYCLNVPLDDGIDD 246

Qy    243 ESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVP 302
          :||  :|||| |||||::|||||||||||||||||||||||| ||||||||||||||
Db    247 DSYQSIFKPIISKVMEMYRPGAVVLQCGADSLSGDRLGCFNLSGKGHAECVKFMRSFNVP 306

Qy    303 LLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKN 362
          ||||||||| |||||||||||||||| |: |||| :||:|||||:|:| || |||||:|
Db    307 LLLLGGGGYAIRNVARCWCYETGVALGHELTDKMPPNEYFEYFGPEYSLFVAASNMENRN 366

Qy    363 SRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDD 422
          : : ||||: ::| |||||||||| |:|| |:|: || |||||| |:| |||||| | |
Db    367 TNKQLEEIKCNILDNLSKLQHAPSVQFEERIPETKLPEPDEDQEDPDERHDPDSDM-VLD 425

Qy    423 DRKPIPS---------RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472
          | ||:           |||| | : ||: | | :    || |   |: ||:|   |:
Db    426 DHKPMGHSARSLIRNIEVKREITESEAKDQHG-KRLTTEHKGPEPMADDPGSSKQAPVS 483


RESULT   4
Q9FML2
ID   Q9FML2      PRELIMINARY;      PRT;   471 AA.
AC   Q9FML2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=98162728; PubMed=9501997;
RA   Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT   Sequence features of the regions of 1,191,918 bp covered by seventeen
RT   physically assigned P1 clones.";
RL   DNA Res. 4:401-414(1997).
DR   EMBL; AB008265; BAB10553.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   471 AA;  52651 MW;  CA16C2640D1B1732 CRC64;


  Query Match             57.6%;  Score 1568.5;  DB 10;  Length 471;
  Best Local Similarity   60.9%;  Pred. No. 7.8e-115;
  Matches  293;  Conservative   64;  Mismatches   93;  Indels   31;  Gaps    5;

Qy      2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
          |  | || ||||| ||:| |||:| :|:|||||||||||||| |:|: || | : :::
Db      4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63

Qy     62 LKPFPARERDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119
           :|  |   |: |||: :|| || |::||:  |    | |:|||||||||||||: ||:
Db     64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123

Qy    120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179
            ||||:| :||||    |||||| |||||||| |||||||||||||| ||||||  :|||
Db    124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183

Qy    180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239
          |:|||:|||||||||| ||||||||||||||:|||||:|:|   ||||:|||||:||
Db    184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243

Qy    240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299
          :|||:  ||:|:: ||||:::| |||||||||||||||||||||:||||:|:::|:||:
Db    244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303

Qy    300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359
          ||||::|||||||||||||||| ||:||| ::|:| :||:|||||||||| || ||
Db    304 NVPLMVLGGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363

Qy    360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDP----- 414
          | |: : :| ||| ||  || | ||||| ||  ||      :|| ::| : |  |
Db    364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRIWSG 421

Qy    415 --DSDMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472
              : | ||| ||                  | |   || |   : | :|  ::   |
Db    422 TATYESDSDDDDKP------------------LHGYSCRG-GATTDRDSTGEDEMDDDN 461

Qy    473 P 473
          |
Db    462 P 462


RESULT   5
Q9FVE5
ID   Q9FVE5      PRELIMINARY;      PRT;   471 AA.
AC   Q9FVE5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE.
GN   RPD3B.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Wu K., Malik K., Tian L., Brown D., Miki B.;
RT   "Functional analysis of RPD3 histone deacetylase homologs in
RT   Arabidopsis thaliana.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF195548; AAG28475.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   471 AA;  52720 MW;  371BF7040E508849 CRC64;


  Query Match             57.4%;  Score 1561.5;  DB 10;  Length 471;
  Best Local Similarity   60.7%;  Pred. No. 2.8e-114;
  Matches  292;  Conservative   64;  Mismatches   94;  Indels   31;  Gaps    5;

Qy      2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
          |  | || ||||| ||:| |||:| :|:|||||||||||||| |:|: || | : :::
Db      4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63

Qy     62 LKPFPARERDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119
           :|  |   |: |||: :|| || |::||:  |    | |:|||||||||||||: ||:
Db     64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123

Qy    120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179
            ||||:| :||||    |||||| |||||||| |||||||||||||| ||||||  :|||
Db    124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183

Qy    180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239
          |:|||:|||||||||| ||||||||||||||:|||||:|:|   ||||:|||||:||
Db    184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243

Qy    240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299
          :|||:  ||:|:: ||||:::| |||||||||||||||||||||:||||:|:::|:||:
Db    244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
```

```
Qy     300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359
           ||||::|||| |||||||||||||||| ||:||| ::|:| :||:|||||||||||| || ||
Db     304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363

Qy     360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDS--- 416
           | |: : :| ||| ||  || | ||||| ||  ||      :|| ::| : |  |
Db     364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXWSG 421

Qy     417 ----DMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472
               : | ||| ||                  | |   }} |   : | :|  ::   |
Db     422 TATYESDSDDDDKP------------------LHGYSCRG-GATTDRDSTGEDEMDDDN 461

Qy     473 P 473
           |
Db     462 P 462


RESULT   6
Q9ZTP8
ID   Q9ZTP8      PRELIMINARY;      PRT;   458 AA.
AC   Q9ZTP8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
GN   HD1B.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. CUZCO;
RA   Pipal A., Wegener S.;
RL   Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF045473; AAD10139.1; -.
DR   Mendel; 39025; Zeama;3043;39025.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   458 AA;  50940 MW;  A62775068225BE79 CRC64;


  Query Match             56.9%;  Score 1548;  DB 10;  Length 458;
  Best Local Similarity   63.3%;  Pred. No. 3e-113;
  Matches  280;  Conservative   66;  Mismatches   80;  Indels   16;  Gaps    3;

Qy       7 SLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFP 66
           | | | |  :|:| |||:| :|:|||||||||||||||| |:|: |||| : ::: :|:|
Db      12 SPAGGEDAHRRRVSYFYEPSIGDYYYGQGHPMKPHRIRMAHSLVVHYGLHRLLELSRPYP 71

Qy      67 ARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSV 126
           | | |: |||:||||:|| | |       | :|||||||||||||: |||  ||||:
Db      72 ASEADIRRFHSDDYVAFLASATGNPGVLDPRAIKRFNVGEDCPVFDGLFPFCQASAGGSI 131

Qy     127 GGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIH 186
           | :|||| |  || :|||||||||| |||||||||||||||||||| | ||||||||:|
Db     132 GAAVKLNRGDADITVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKFHRRVLYVDIDVH 191

Qy     187 HGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYH 246
           |||||||||: |:|||||||||:||:|||||| |:|   ||:|:|||| |||| ::
Db     192 HGDGVEEAFFTTNRVMTVSFHKYGDFFPGTGHITDVGAAEGKHYALNVPLSDGIDDTTFR 251

Qy     247 LLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLL 306
            ||: |: |||||:|  ||||||||||:|||||||||:||||:|::|:||:|||:::|
Db     252 GLFQCIIKKVMEVYQPDVVVLQCGADSLAGDRLGCFNLSVKGHADCLRFLRSYNVPMMVL 311

Qy     307 GGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQM 366
           |||||||||||||||| ||:||| ::|:| ::|||||||||||: } ::|| |: :
Db     312 GGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNDYYEYFGPDYTLHIQPKSVENLNTTKD 371

Qy     367 LEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKP 426
```

```
           || |:| :| ||||::| ||  | :|| | | ||  | :|| ||            | |
Db     372 LENIKNMILENLSKIEHVPSTQFHDRPSDPEAPE--EKEEDMDK------------RPP 416

Qy     427 IPSRV-KREAVEPDTKDKDGLK 447
             ||:    | : ||:| | ||
Db     417 QRSRLWSGGAYDSDTEDPDSLK 438


RESULT   7
O77213
ID   O77213      PRELIMINARY;      PRT;   521 AA.
AC   O77213;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE.
GN   RPD3 OR CG7471.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mottus R.C., Sobel R.E., Grigliatti T.A.;
RT   "DmHDAC1.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF086715; AAC61494.1; -.
DR   FlyBase; FBgn0015805; Rpd3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   521 AA;  58344 MW;  B0F6503D42A1BA32 CRC64;


  Query Match             55.0%;  Score 1496.5;  DB 5;  Length 521;
  Best Local Similarity   56.3%;  Pred. No. 4e-109;
  Matches  268;  Conservative   80;  Mismatches  103;  Indels   25;  Gaps    4;

Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           |::|||:|| ::|||||||||||||||||||| || :||| : |:: :|  |   :: :|
Db       6 KKRVCYYYDSDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKATADEMTKF 65

Qy      76 HADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHG 135
           |:|:|| ||||| |:   :  :|::|||||||||||||||| |||  |||||  :||||
Db      66 HSDEYVRFLRSIRPDNMSEYNKQMQRFNVGEDCPVFDGLYEFCQLSAGGSVAAAVKLNKQ 125

Qy     136 LCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAF 195
             :| ||| |||||||| |||||||||||| |||||| |:||||:|||:||||||||||
Db     126 ASEICINWGGGLHHAKKSEASGFCYVNDIVLGILELLKYHQRVLYIDIDVHHGDGVEEAF 185

Qy     196 YATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           | |||||||||||:|:|||||| ::||| | ||||::|:|| ||:||::|  :| ||: |
Db     186 YTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNIPLRDGMDDDAYESIFVPIISK 245

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           ||| |:| ||||||||||:|||||||||::||| :||:|:: :|:| |::|||||||||
Db     246 VMETFQPAAVVLQCGADSLTGDRLGCFNLTVKGHGKCVEFVKKYNLPFLMVGGGGYTIRN 305

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLL 375
           |:||| ||| ||| ||: :::| ::|:|||||: ||::|||| |:|: : ||:|:| |
Db     306 VSRCWTYETSVALAVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTSEYLEKIKNRLF 365

Qy     376 HNLSKLQHAPSVPFQERPPDTETPEVDED----------QEDGDKRWDPDSD-MDVDD-- 422
           ||  | ||| | |  | | |    | |::          | | |||  |:::  | :|
Db     366 ENLRMLPHAPGVQIQAIPEDAINDESDDEDKVDKDDRLPQSDKDKRIVPENEYSDSEDEG 425

Qy     423 -----DRKPIPSRVKREAVEPDTKDKDG-------LKGIMERGKGCEVEVDESGST 466
                | :    : ||  :: ||            :|   |:| | : |   : :|
Db     426 EGGRRDNRSYKGQRKRPRLDKDTNSNKASSETSSEIKDEKEKGDGADGEESTASNT 481


RESULT   8
```

```
Q9VZA1
ID   Q9VZA1      PRELIMINARY;      PRT;   521 AA.
AC   Q9VZA1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   RPD3 PROTEIN.
GN   RPD3 OR CG7471.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003482; AAF47924.1; -.
DR   FlyBase; FBgn0015805; Rpd3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   521 AA;  58330 MW;  B0F6503D42A1BCE9 CRC64;
```

```
  Query Match             55.0%;  Score 1496.5;  DB 5;  Length 521;
  Best Local Similarity   56.3%;  Pred. No. 4e-109;
  Matches  268;  Conservative   80;  Mismatches  103;  Indels   25;  Gaps    4;

Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           |::|||:|| ::|||||||||||||||||||||| || :||| : |:: :|  |   :: :|
Db       6 KKRVCYYYDSDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKATADEMTKF 65

Qy      76 HADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHG 135
           |:|:|| ||||| |:   :  :|::|||||||||||||||| |||  |||||  :||||
Db      66 HSDEYVRFLRSIRPDNMSEYNKQMQRFNVGEDCPVFDGLYEFCQLSAGGSVAAAVKLNKQ 125

Qy     136 LCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAF 195
             :| ||| |||||||| |||||||||||||| |||||| |:||||:|||:|||||||||
Db     126 ASEICINWGGGLHHAKKSEASGFCYVNDIVLGILELLKYHQRVLYIDIDVHHGDGVEEAF 185

Qy     196 YATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           | |||||||||||:|:|||||| ::||| | ||||::|:|| ||:||::|  :| ||: |
Db     186 YTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNIPLRDGMDDDAYESIFVPIISK 245

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           ||| |:| ||||||||||:|||||||||::||| :||:|:: :|:| |::|||||||||
Db     246 VMETFQPAAVVLQCGADSLTGDRLGCFNLTVKGHGKCVEFVKKYNLPFLMVGGGGYTIRN 305

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLL 375
           |:||| ||| ||| ||: :::| ::|:|||||: ||::|||| |:|: : ||:|:| |
Db     306 VSRCWTYETSVALAVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTSEYLEKIKNRLF 365

Qy     376 HNLSKLQHAPSVPFQERPPDTETPEVDED----------QEDGDKRWDPDSD-MDVDD-- 422
           ||  | ||| |  |  | |    | |::          | | |||  |::: | :|
Db     366 ENLRMLPHAPGVQIQAIPEDAINDESDDEDKVDKDDRLPQSDKDKRIVPENEYSDSEDEG 425

Qy     423 -----DRKPIPSRVKREAVEPDTKDKDG-------LKGIMERGKGCEVEVDESGST 466
                | :   : ||  :: ||           :|   |:| | : |   : :|
Db     426 EGGRRDNRSYKGQRKRPRLDKDTNSNKASSETSSEIKDEKEKGDGADGEESTASNT 481
```

```
RESULT   9
Q99PA0
ID   Q99PA0      PRELIMINARY;      PRT;   428 AA.
AC   Q99PA0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE 3.
GN   HDAC3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=TESTIS;
RA   Wilquet V., Chavez M., Korbers R., Geerts A.;
RT   "Expression pattern of rat histone deacetylases.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF321131; AAK11184.1; -.
SQ   SEQUENCE   428 AA;  48829 MW;  7CB448CFE33041FE CRC64;
```

```
  Query Match             50.5%;  Score 1374;  DB 11;  Length 428;
  Best Local Similarity   56.6%;  Pred. No. 1.2e-99;
  Matches  241;  Conservative   73;  Mismatches   90;  Indels   22;  Gaps    3;

Qy      17 RKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFH 76
           : | |||||:|||::|| |||||||||: :|||:|: |||| : | | ||: | : |:|||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      77 ADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
           ::||: ||: ::|   |   : |  ||||:||||| ||: ||  | | |: |: :||: :
Db      63 SEDYIDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNNKI 122

Qy     137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
           |||||||||||||||| ||||||||||||: |||||| | ||||:|||||||||:||||
Db     123 CDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEAFY 182

Qy     197 ATDRVMTVSFHKFGDY-FPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           |||||||||||:|:|  ||||| : ::|  ||:|| ||:|| |||||:||  ||:|:: :
Db     183 LTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVISQ 242

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           |:: ::|  :|||||||||  ||||||||||:|| |||:::|||:|||:|||||||:||
Db     243 VVDFYQPXCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN 302
```

```
Qy    316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRNDL 374
          ||||| ||| : :   : :::|  ||:||| ||:|||    : :|||:|||| |::||  :
Db    303 VARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQTI 362

Qy    375 LHNLSKLQHAPSVPFQERPPDTET-------------------PEVDEDQEDGDKRWDP 414
            ||  | |||||   : | |  |                    ||   :  |||   |
Db    363 FENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDNDK 422

Qy    415 DSDMDV 420
          :||:::
Db    423 ESDVEI 428


RESULT  10
Q9P4F5
ID   Q9P4F5      PRELIMINARY;      PRT;   687 AA.
AC   Q9P4F5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE RPD3A.
GN   RPD3A.
OS   Emericella nidulans (Aspergillus nidulans).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=5072;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20461766; PubMed=11004483;
RA   Graessle S., Dangl M., Haas H., Mair K., Trojer P., Brandtner E.M.,
RA   Walton J.D., Loidl P., Brosch G.;
RT   "Characterization of two putative histone deacetylase genes from
RT   Aspergillus nidulans.";
RL   Biochim. Biophys. Acta 1492:120-126(2000).
DR   EMBL; AF163862; AAF80489.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   687 AA;  75431 MW;  CFC7C65352366451 CRC64;


  Query Match             50.0%;  Score 1361.5;  DB 3;  Length 687;
  Best Local Similarity   49.2%;  Pred. No. 2.3e-98;
  Matches  261;  Conservative   89;  Mismatches  136;  Indels   45;  Gaps    9;

Qy     10 SGPDGV-------------KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLL 56
          ||| |                ::| |||| :|||| |  |||||||||||||:|: :| |
Db      6 SGPAGPPLDPIDLNVSGDRSKRVAYFYDSDVGNYAYVSGHPMKPHRIRMTHSLVMNYSLY 65

Qy     57 QHMQVLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYS 116
          : |:: :  || : :: :|| |:|: ||  :||:      ::  ::||||:||||||||:
Db     66 KKMEIYRAKPASKFEMTQFHTDEYIDFLSKVTPDNMDAFAKEQSKYNVGDDCPVFDGLFE 125

Qy    117 FCQTYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHE 176
          ||   ||||: |: :||   ||||:|||||||||||| |||||||||||| |||||:  :
Db    126 FCGISAGGSMEGAARLNRNKCDIAVNWAGGLHHAKKSEASGFCYVNDIVLGILELLRFKQ 185

Qy    177 RVLYVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPL 236
          ||||||||:|||||||||||| ||||||||||:|:|||||| ::||| | ||||::| ||
Db    186 RVLYVDIDVHHGDGVEEAFYTTDRVMTVSFHKYGEYFPGTGELRDIGVGQGKYYAVNFPL 245

Qy    237 DDGIDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFM 296
           ||||| ||  :|:|::  ||| :|| |||||| ||||||||||||||::||| |||:: 
Db    246 RDGIDDVSYKSIFEPVIKSVMEWYRPEAVVLQCGGDSLSGDRLGCFNLSMRGHANCVKYV 305

Qy    297 RSFNVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPS 356
          :|||:| |::|||||:||||| | :|||: :|  :  ::| ::||||| ||| | | ||
Db    306 KSFNLPTLIVGGGGYTMRNVARTWAFETGILVGDNLGSELPYNDYYEYFAPDYELDVRPS 365

Qy    357 NMENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDS 416
          ||:| |:|: |::||  :: || :   ||||   : | :   | ||   :: :   | |
Db    366 NMDNANTREYLDKIRTQVVENLKRTAFAPSVQMTDVPRE---PLVDGMDDEAEAALD-DL 421
```

```
Qy    417 DMDVDDDRKPIPSRVKREAVEP----DTKDKD--GLKGI------MERGKGCEVEVD--- 461
          | | : |::    |  :   :|    |::|:|     |:       ::|       :|
Db    422 DEDENKDKRFTKRRFDQYVEKPGELSDSEDEDENAANGVTRKPAHLKRRNQANYRLDLAD 481

Qy    462 ---ESGSTKVTGVNPVGVEE----ASVKM------EEEGTNKGGAEQAFPP 499
             |||       : |  ||      ||:      | :   :| :  | ||
Db    482 SGVESGMATPQDASSVADEEMDTGTDVKITEAPGPEPDSEAQGTSSAAEPP 532


RESULT  11
Q9JM08
ID   Q9JM08      PRELIMINARY;      PRT;   428 AA.
AC   Q9JM08;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE 3.
GN   HDAC3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=20391214; PubMed=10542131;
RA   Dangond F., Foerznler D., Weremowicz S., Morton C.C., Beier D.R.,
RA   Gullans S.R.;
RT   "Cloning and expression of a murine histone deacetylase 3 (mHdac3)
RT   cDNA and mapping to a region of conserved synteny between murine
RT   chromosome 18 and human chromosome 5.";
RL   Mol. Cell Biol. Res. Commun. 2:91-96(1999).
DR   EMBL; AF098295; AAF28798.1; -.
DR   MGD; MGI:1343091; Hdac3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   428 AA;  48890 MW;  995EFFB8275688D4 CRC64;


  Query Match             49.9%;  Score 1359;  DB 11;  Length 428;
  Best Local Similarity   56.1%;  Pred. No. 1.8e-98;
  Matches  239;  Conservative   73;  Mismatches   92;  Indels   22;  Gaps    3;

Qy     17 RKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFH 76
          : | |||||:|||::|| |||||||||: :||:|: |||| : | | ||: | : |:|||
Db      3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy     77 ADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
          ::||: ||: ::|   |   : |  ||||:||||| ||: ||  | | |: |: :||: :
Db     63 SEDYIDFLQRVSPTNMQGFTKSLDAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNNKI 122

Qy    137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
          |||||||| |||||| ||||||||||||: |||||| | ||||:||||||||||:||||
Db    123 CDIAINWARGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEAFY 182

Qy    197 ATDRVMTVSFHKFGDY-FPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
          ||||||||||:|:| ||| | : ::|  ||:|| ||||| |||||:||  ||:|:: :
Db    183 LTDRVMTVSFHKYGNYFFPGAGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVISQ 242

Qy    256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
          |:: ::|  :|||||||||  ||||||||||:|| |||::::|||:|||:||||||:||
Db    243 VVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN 302

Qy    316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRNDL 374
          ||||| ||| : :   : :::|  ||:||| ||:|||    : :|||:|||| |::||  :
Db    303 VARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQTI 362

Qy    375 LHNLSKLQHAPSVPFQERPPDTET-------------------PEVDEDQEDGDKRWDP 414
            ||  | |||||   : | |  |                    ||   :  |||   |
Db    363 FENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDNDK 422
```

```
Qy     415 DSDMDV 420
           :|||:::
Db     423 ESDVEI 428


RESULT  12
Q9JLX5
ID   Q9JLX5      PRELIMINARY;      PRT;   428 AA.
AC   Q9JLX5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE-3.
GN   HDAC3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tsai S.C., Kahle E., Seto E.;
RT   "Isolation and characterization of the mouse histone deacetylase-3
RT   gene.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF125536; AAF36425.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   428 AA;  48737 MW;  EDE9E2B43045119A CRC64;


  Query Match             49.9%;  Score 1358;  DB 11;  Length 428;
  Best Local Similarity   56.6%;  Pred. No. 2.2e-98;
  Matches  240;  Conservative   74;  Mismatches   92;  Indels   18;  Gaps    4;

Qy      17 RKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFH 76
           : | |||||:|||::|| ||||||||: :||:|: |||| : | | ||: | : |:|||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      77 ADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
           ::|||: ||: ::|   |   : |  ||||:||||| ||: ||  | | |: |: :|||: :
Db      63 SEDYIDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNNKI 122

Qy     137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
           |||||||||||||||| ||||||||||||: ||| || | ||||:|||||||||:|||| 
Db     123 CDIAINWAGGLHHAKKFEASGFCYVNDIVIGILEPLKYHPRVLYIDIDIHHGDGVQEAFY 182

Qy     197 ATDRVMTVSFHKFGDY-FPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           |||||||||||:|:| ||||| : ::|  ||:|| ||||| |||||:||  ||:|:: :
Db     183 LTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVISQ 242

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           |:: ::|  :|||||||||  ||||||||||:|| |||:::|||:|||:|||||||:||
Db     243 VVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN 302

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRNDL 374
           ||||| ||| : :   : :::|  ||:||| ||:|||    : :||:|||| |::|   :
Db     303 VARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQICQTI 362

Qy     375 LHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKR--------------WDPDSDMD 419
             ||  | |||||   : | |  | :   ||: ||::|                :| | | |
Db     363 FENLKMLNHAPSVQIHDVPADLLTYD-RTDEADGEERGPEENYSRPEAPNEFYDGDHDND 421

Qy     420 VDDD 423
           : |
Db     422 KESD 425


RESULT  13
O59702
ID   O59702         PRELIMINARY;   PRT;   405 AA.
AC   O59702;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
GN   SPBC36.05C OR CLR6.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972H-;
RA   Lyne M., Wood V., Rajandream M.A., Barrell B.G., Hilbert H.,
RA   Moestl D., Duesterhoeft A.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SP972;
RX   MEDLINE=98429513; PubMed=9755190;
RA   Grewal S.I., Bonaduce M.J., Klar A.J.;
RT   "Histone deacetylase homologs regulate epigenetic inheritance of
RT   transcriptional silencing and chromosome segregation in fission
RT   yeast.";
RL   Genetics 150:563-576(1998).
DR   EMBL; AL023589; CAA19053.1; -.
DR   EMBL; AF064206; AAD05211.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   405 AA;  46112 MW;  8EDEA43D5839E367 CRC64;


  Query Match             49.1%;  Score 1336;  DB 3;  Length 405;
  Best Local Similarity   61.7%;  Pred. No. 1.1e-96;
  Matches  235;  Conservative   62;  Mismatches   82;  Indels    2;  Gaps    2;

Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           +:|| |||| :||||:|| |||||||:|| | |: :| | : : |: |  |   |: |
Db       5 KKKVSYFYDEDVGNYHYGPQHPMKPHRVRMVHNLVVNYNLYEKLNVITPVRATRNDMTRC 64

Qy      76 HADDYVSFLRSITPETQQD-QIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNH 134
           | |:|: ||  :||:| :  |  ||| ||||:||||||||| ||   ||||:| : :||
Db      65 HTDEYIEFLWRVTPDTMEKFQPHQLK-FNVGDDCPVFDGLYEFCSISAGGSIGAAQELNS 123

Qy     135 GLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEA 194
           |  :|||||||||||||| |||||||||| || ||||| |:||||:|||:||||||||
Db     124 GNAEIAINWAGGLHHAKKREASGFCYVNDIALAALELLKYHQRVLYIDIDVHHGDGVEEF 183

Qy     195 FYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMG 254
           || |||||| ||||||:||||||||:| | |:|| |::|||| ||||||||  :|||::
Db     184 FYTTDRVMTCSFHKFGEYFPGTGHIKDTGIGTGKNYAVNVPLRDGIDDESYESVFKPVIS 243

Qy     255 KVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIR 314
           :|: ||| ||:|||| |||:|||||||||||:||| || |::|||:|:: :|||||:|
Db     244 HIMQWFRPEAVILQCGTDSLAGDRLGCFNLSMKGHSMCVDFVKSFNLPMICVGGGGYTVR 303

Qy     315 NVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDL 374
           |||| | ||||:  | |::: :| ::| :|:|||| |:|  :|||| |:|| |: | :::
Db     304 NVARVWTYETGLLAGEELDENLPYNDYLQYYGPDYKLNVLSNNMENHNTRQYLDSITSEI 363

Qy     375 LHNLSKLQHAPSVPFQERPPD 395
           : ||  |  ||||   : | |
Db     364 IENLRNLSFAPSVQMHKTPGD 384


RESULT  14
O62339
ID   O62339      PRELIMINARY;   PRT;   465 AA.
AC   O62339; O62343;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   R06C1.1 PROTEIN.
```

```
GN   R06C1.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kershaw J.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z81108; CAB03240.1; -.
DR   EMBL; Z81106; CAB03240.1; JOINED.
DR   EMBL; Z81106; CAB03224.1; -.
DR   EMBL; Z81108; CAB03224.1; JOINED.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   465 AA;  52709 MW;  325E71D5197012C5 CRC64;
```

```
  Query Match           49.0%;  Score 1334;  DB 5;  Length 465;
  Best Local Similarity  55.9%;  Pred. No. 1.9e-96;
  Matches  246;  Conservative   74;  Mismatches  106;  Indels   14;  Gaps    5;

Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           | :| |:|| : ||:|||||||||||||:||||:|: :||| : : |::|  |   :: |:
Db       7 KSRVSYYYDGDFGNFYYGQGHPMKPHRVRMTHSLIVNYGLYRKLNVMRPARASFSEITRY 66

Qy      76 HADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHG 135
           |:|||::|||::  :      |: ||:||||||||||:| |||   |||:  : :||
Db      67 HSDDYINFLRNVKSDNMSTFTDQMARFSVGEDCPVFDGMYEFCQLSCGGSLAAAARLNRQ 126

Qy     136 LCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAF 195
             :||||| |||||||| |||||| ||||||||||| |:||||:|||:|||||||||
Db     127 ESEIAINWMGGLHHAKKSEASGFCYSNDIVLAILELLKHHKRVLYIDIDVHHGDGVEEAF 186

Qy     196 YATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           | |||||||||||  |:|||||| ::|:| |||||:|||||  ||:|| :|  :|: |||:
Db     187 YTTDRVMTVSFHKHGEYFPGTGDLKDVGAGSGKYYALNVPLRDGVDDVTYERIFRTIMGE 246

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           ||  |:| |||||||||||:|||||  |||:  || :||::|:|||||||:|||||||||
Db     247 VMARFQPEAVVLQCGADSLAGDRLGVFNLTTYGHGKCVEYMKSFNVPLLLVGGGGYTIRN 306

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAP-SNMENKNSRQMLEEIRNDL 374
           |:||| ||| :||  || | :| |:|::|| ||| ||: | : : | |: : :::    |
Db     307 VSRCWLYETAIALNQEVSDDLPLHDYFDYFIPDYKLHIKPLAALSNFNTPEFIDQTIVAL 366

Qy     375 LHNLSKLQHAPSVPFQ--ERPPDTETPEVDE----DQEDGDKRWDP-DSDMDVDDDRKPI 427
           | || :| | |||  |      |:     ||    | :: | |    : |: |:|
Db     367 LENLKQLPHVPSVQMQSISTSCDSIVKTFDEKLIRDHQNDDVRVTQFEEDVQVED----- 421

Qy     428 PSRVKREAVEPDTKDKDGLK 447
           |    :  ||:||:   :|
Db     422 -SAEFYDGQEPETKNIQSMK 440
```

RESULT 15

```
Q9C1C6
ID   Q9C1C6      PRELIMINARY;      PRT;   648 AA.
AC   Q9C1C6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE 2.
GN   HDC2.
OS   Cochliobolus carbonum (Bipolaris zeicola).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC   Pleosporales; Pleosporaceae; Cochliobolus.
OX   NCBI_TaxID=5017;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wegener S., Walton J.D.;
RT   "HDC2, a histone deacetylase gene related to RPD3, from the
RT   filamentous fungus Cochliobolus carbonum.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF349677; AAK35180.1; -.
SQ   SEQUENCE   648 AA;  70940 MW;  242147C293D4EAD3 CRC64;
```

```
  Query Match           48.8%;  Score 1329.5;  DB 3;  Length 648;
  Best Local Similarity  46.0%;  Pred. No. 6.7e-96;
  Matches  260;  Conservative   83;  Mismatches  145;  Indels   77;  Gaps    9;

Qy       5 GNSLASGPDGV----------KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYG 54
           || :| || |            :|| |||| :|||| |  ||||||||||| |:|: :||
Db       4 GNDMA-GPTGAVHINNGTASQPKKVAYFYDSDVGNYAYVAGHPMKPHRIRMAHSLIMNYG 62

Qy      55 LLQHMQVLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGL 114
           |   |:: :  || : :: :|| |:|: ||  :||:  :   |:  ::|||:|||||||
Db      63 LYTKMEIYRAKPASKYEMTQFHTDEYIEFLHKVTPDNMEQFTREQSKYNVGDDCPVFDGL 122

Qy     115 YSFCQTYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQ 174
           : ||   |||:: |: :|| | ||:|:|||||||||||| ||||||||||||| |:|||:
Db     123 FEFCGISAGGTMEGAARLNRGKCDVAVNWAGGLHHAKKSEASGFCYVNDIVLGIIELLRY 182

Qy     175 HERVLYVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNV 234
           :||||:|||:|||||||||| |||||||||||:|:|||||| ::||| |||| ||:|
Db     183 KQRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKYGEYFPGTGELRDIGVGSGKNYSVNF 242

Qy     235 PLDDGIDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVK 294
           || ||| ||:|  :|:|::  ||  : | |:||||| |||||||||||||: ||| ||:
Db     243 PLRDGITDETYRNIFEPVIEAVMTYYGPEAIVLQCGGDSLSGDRLGCFNLSMDGHANCVR 302

Qy     295 FMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVA 354
           :::|| ||:::|||||||:|||| | |||  :  ::  ::| ::|||| ||| | |
Db     303 YVKSFGVPVIVLGGGGYTMRNVARTWAYETGELVSQKMSKQLPFNDYYEYFAPDYELDVR 362

Qy     355 PSNMENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETP---EVDEDQEDGDKR 411
           |||||| ||   | :|:: :: |: :   |||      ||  |     :| | ||: :
Db     363 PSNMENANSHDYLHKIKSAVIENIRRTGR-PSVEAFTTIPDVPTALGRAMDSDAEDEEDD 421

Qy     412 WDPDSDMDV----------------------DDDRK--------------------P1PS 429
           | | : ||                      |:| |                     | |:
Db     422 MDADENRDVRMTQRQRDKQIEHDGELYDASDDEDYKNNLGVRAQPGVKKRRNITDFPNPN 481

Qy     430 RVKREAVE-----------PDTKDKDGLKGIMERGKGCEVEVDESG----STKVTGVNPV 474
           :|::                :  | |   :     | |||: |       |    |
Db     482 AAPADALDLMEDINGGSAVSTRQPSAGAKSRTQTPAAAEQEVDKDGDVDMDAPVPESAPE 541

Qy     475 GVEEASVKMEEEGTNKGGAEQAFPP 499
           | ||:::   |      |  | ||
Db     542 PVPEAALETAPE-----TAPAAAPP 561
```

Search completed: April 26, 2002, 17:57:27
Job time: 355 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:52:17 ; Search time 17.07 Seconds
(without alignments)
1076.103 Million cell updates/sec

Title: US-09-645-337-2
Perfect score: 2722
Sequence: 1 MDTGGNSLASGPDGVKRKVC.....KMEEEGTNKGGAEQAFPPKT 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2722	100.0	501	1	HDAC_ARATH	O22446 arabidopsis
2	1986.5	73.0	513	1	HDAC_MAIZE	P56521 zea mays (m
3	1554	57.1	482	1	HDA1_MOUSE	O09106 mus musculu
4	1544	56.7	482	1	HDA1_HUMAN	Q13547 homo sapien
5	1538	56.5	480	1	HDA1_CHICK	P56517 gallus gall
6	1521	55.9	480	1	HD12_XENLA	O42227 xenopus lae
7	1520	55.8	480	1	HD11_XENLA	Q91695 xenopus lae
8	1513.5	55.6	488	1	HDA2_MOUSE	P70288 mus musculu
9	1507	55.4	576	1	HDA1_STRPU	P56518 strongyloce
10	1494	54.9	488	1	HDA2_HUMAN	Q92769 homo sapien
11	1478	54.3	520	1	HDAC_DROME	Q94517 drosophila
12	1465.5	53.8	488	1	HDA2_CHICK	P56519 gallus gall
13	1392	51.1	461	1	HDA1_CAEEL	O17695 caenorhabdi
14	1376	50.6	428	1	HDA3_CHICK	P56520 gallus gall
15	1373	50.4	428	1	HDA3_HUMAN	O15379 homo sapien
16	1366	50.2	424	1	HDA3_MOUSE	O88895 mus musculu
17	1341	49.3	433	1	RPD3_YEAST	P32561 saccharomyc
18	1167.5	42.9	434	1	PHD1_SCHPO	O13298 schizosacch
19	1077	39.6	507	1	HDA2_CAEEL	Q09440 caenorhabdi
20	1066.5	39.2	452	1	HOS2_YEAST	P53096 saccharomyc
21	471.5	17.3	470	1	HOS1_YEAST	Q12214 saccharomyc
22	443.5	16.3	387	1	ACUC_BACSU	P39067 bacillus su
23	386.5	14.2	385	1	ACUC_STAXY	Q56195 staphylococ
24	352	12.9	359	1	Y130_ARCFU	O30107 archaeoglob
25	320.5	11.8	706	1	HDA1_YEAST	P53973 saccharomyc
26	298.5	11.0	1215	1	HDA6_HUMAN	Q9ubn7 homo sapien
27	278	10.2	687	1	HDA1_SCHPO	P56523 schizosacch
28	275.5	10.1	1080	1	HDA4_CHICK	P83038 gallus gall
29	275.5	10.1	1149	1	HDA6_MOUSE	Q9z2v5 mus musculu
30	274	10.1	1113	1	HDA5_MOUSE	Q9z2v6 mus musculu
31	266.5	9.8	1122	1	HDA5_HUMAN	Q9uql6 homo sapien
32	263	9.7	1084	1	HDA4_HUMAN	P56524 homo sapien
33	261.5	9.6	310	1	YGLA_SYNP2	P28606 synechococc
34	254	9.3	798	1	YLFN_CAEEL	Q20296 caenorhabdi
35	238	8.7	331	1	YB94_METTH	O27262 methanobact
36	230	8.4	343	1	Y535_METJA	Q57955 methanococc
37	208	7.6	304	1	Y245_SYNY3	P72702 synechocyst
38	185	6.8	341	1	APHA_MYCRA	Q48935 mycoplana r
39	127.5	4.7	697	1	HOS3_YEAST	Q02959 saccharomyc
40	117	4.3	1862	1	ANK1_MOUSE	Q02357 mus musculu
41	112	4.1	589	1	HIBN_XENLA	P06180 xenopus lae
42	111.5	4.1	488	1	CYL2_BOVIN	Q28092 bos taurus
43	108	4.0	294	1	NPM_CHICK	P16039 gallus gall
44	108	4.0	2009	1	SEC7_YEAST	P11075 saccharomyc
45	107	3.9	958	1	YGX7_YEAST	P53076 saccharomyc

ALIGNMENTS

RESULT 1
HDAC_ARATH
ID HDAC_ARATH STANDARD; PRT; 501 AA.
AC O22446;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE (HD).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Tomihama T., Shoji K., Hanyu H., Okano T.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC ----------

CC ----------
DR EMBL; AF014824; AAB66486.1; -.
DR InterPro; IPR000286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
DR PRINTS; PR01271; HISDACETLASE.
KW Hydrolase; Nuclear protein.
SQ SEQUENCE 501 AA; 56037 MW; C50AF5624958D6C2 CRC64;

Query Match 100.0%; Score 2722; DB 1; Length 501;
Best Local Similarity 100.0%; Pred. No. 1.2e-203;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
```

```
Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360

Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420

Qy     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480

Qy     481 VKMEEEGTNKGGAEQAFPPKT 501
           |||||||||||||||||||||
Db     481 VKMEEEGTNKGGAEQAFPPKT 501
```

```
RESULT   2
HDAC_MAIZE
ID   HDAC_MAIZE     STANDARD;      PRT;   513 AA.
AC   P56521;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE (RPD3 HOMOLOG).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. W22;
RA   Rossi V., Hartings H., Motto M.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF035815; AAC50038.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   513 AA;  57546 MW;  C45387CF3A38906F CRC64;
```

```
  Query Match           73.0%;  Score 1986.5;  DB 1;  Length 513;
  Best Local Similarity  75.2%;  Pred. No. 1.3e-146;
  Matches  377;  Conservative   44;  Mismatches   61;  Indels   19;  Gaps    7;

Qy       3 TGGNSLAS-GPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
             :|||||  | |||| ||:|||||||:|||||||||||||||||||||:||| |||| |||
Db       8 SGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPMKPHRIRMTHSLLARYGLLNQMQV 67

Qy      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTY 121
             :| |||||:||||||::|:::||||:|||||||||| |||||||:||| ||||||||||
Db      68 YRPNPARERELCRFHAEEYINFLRSVTPETQQDQIRLLKRFNVGEECPVLDGLYSFCQTY 127

Qy     122 AGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYV 181
           || ||||:|| ||| |||||||:|||||||||||||||||||||||||| |||||||
Db     128 AGASVGGAVKFNHG-HDIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKHHERVLYV 186

Qy     182 DIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGID 241
           |||||||||||||| |||||||||||||||||||| |:|||: ||||||||||||||||
Db     187 DIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGHSKGKYYSLNVPLDDGID 246

Qy     242 DESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNV 301
           ||||  |||||||||||:|||||||||||||||||||||||||||||||||||::||||||
Db     247 DESYQSLFKPIMGKVMEVFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYMRSFNV 306

Qy     302 PLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENK 361
           ||||||||||||||||||||||||||||| | ||||| :||||||||||||||||||||||
Db     307 PLLLLGGGGYTIRNVARCWCYETGVALGQEPEDKMPVNEYYEYFGPDYTLHVAPSNMENK 366

Qy     362 NSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVD 421
           |:|| |::||:     ||||:||||| |||| |||| || ||||:| |:| ||||||:||
Db     367 NTRQQLDDIRS----KLSKLRHAPSVHFQERVPDTEIPEQDEDQDDPDERHDPDSDMEVD 422

Qy     422 DD--------RKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK---VTG 470
           |         |  :  ::|||  |  |: :|| : :    :| |   :: ||:|
Db     423 DHKAVEESSRRSILGIKIKREFGENATRVQDGGR-VASEHRGLEPMAEDIGSSKQAPQAD 481

Qy     471 VNPVGVEEAS-VKMEEEGTNK 490
           : : ::| | || | | : |
Db     482 ASAMAIDEPSNVKNEPESSTK 502
```

```
RESULT   3
HDA1_MOUSE
ID   HDA1_MOUSE     STANDARD;      PRT;   482 AA.
AC   O09106; P97476;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=97415582; PubMed=9271381;
RA   Bartl S., Taplick J., Lagger G., Khier H., Kuchler K., Seiser C.;
RT   "Identification of mouse histone deacetylase 1 as a growth factor-
RT   inducible gene.";
RL   Mol. Cell. Biol. 17:5033-5043(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Johnson C.A.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- FUNCTION: ACTS AS A TRANSCRIPTIONAL REPRESSOR.
CC   -!- SUBUNIT: FORMS A COMPLEX WITH RBAP48 (BY SIMILARITY). ALSO FORMS A
```

```
CC       COMPLEX WITH SIN3 AND SAP18 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHER LEVELS IN THYMUS
CC       AND TESTIS AND LOWER LEVELS IN LIVER.
CC   -!- INDUCTION: BY INTERLEUKIN-2.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X98207; CAA66870.1; -.
DR   EMBL; U80780; AAB68398.1; -.
DR   MGD; MGI:108086; Hdac1.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   482 AA;  55075 MW;  7F64D3C17F5E4844 CRC64;


  Query Match             57.1%;  Score 1554;  DB 1;  Length 482;
  Best Local Similarity   60.0%;  Pred. No. 4.3e-113;
  Matches  294;  Conservative   64;  Mismatches   98;  Indels   34;  Gaps    7;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | ||||||:|| :|||||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 GTKRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::|||||||||||||: |||   |||| :||||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVKLN 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
           |||:|||||||||| |||||||||||||||||| |:|||||:|||||||||||
Db     126 KQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGVEE 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| ||||||||||:|:|||||| ::||| | ||||::| || ||||||||  :|||:|
Db     186 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVM 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           ||||:|:| |||||||:|||||||||||||:|||||:||:|:::|||:|:|:|||||||
Db     246 SKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGYTI 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           ||||||| ||| |||  |: :::| ::|:|||||: ||::|||| |:|: : ||:|:
Db     306 RNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIKQR 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD----------PDSDM 418
           |  ||  | ||| |  |  | |    |  |||:|| |||              |||
Db     366 LFENLRMLPHAPGVQMQAIPEDAIPEESGDEDEEDPDKRISICSSDKRIACEEEFSDSDE 425

Qy     419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
           : :  ||   :     ||| |    | |:||      |  |    || |   ||
Db     426 EGEGGRKNSSNFKKAKRVKTE----DEKEKD-----PEEKK----EVTEEEKTKEEKPEA 472

Qy     474 VGVEEASVKM 483
           ||:|  ||:
Db     473 KGVKE-EVKL 481


RESULT   4
HDA1_HUMAN
ID   HDA1_HUMAN     STANDARD;      PRT;   482 AA.
AC   Q13547; Q92534;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1 OR RPD3L1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-cell;
RX   MEDLINE=96185499; PubMed=8602529;
RA   Taunton J., Hassig C.A., Schreiber S.L.;
RT   "A mammalian histone deacetylase related to the yeast transcriptional
RT   regulator Rpd3p.";
RL   Science 272:408-411(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal lung;
RX   MEDLINE=96244606; PubMed=8646880;
RA   Furukawa Y., Kawakami T., Sudo K., Inazawa J., Matsumine A.,
RA   Akiyama T., Nakamura Y.;
RT   "Isolation and mapping of a human gene (RPD3L1) that is homologous to
RT   RPD3, a transcription factor in Saccharomyces cerevisiae.";
RL   Cytogenet. Cell Genet. 73:130-133(1996).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS.
CC   -!- SUBUNIT: FORMS A COMPLEX WITH RBAP48 AND ALSO WITH THE HISTONE
CC       ACETYLTRANSFERASE P/CAF AND THE ADAPTATOR PROTEIN P300. ALSO FORMS
CC       A COMPLEX WITH SIN3 AND SAP18. INTERACTS WITH TGIF.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUS, WITH HIGHER LEVELS IN HEART,
CC       PANCREAS AND TESTIS, AND LOWER LEVELS IN KIDNEY AND BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U50079; AAC50475.1; -.
DR   EMBL; D50405; BAA08909.1; -.
DR   MIM; 601241; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   CONFLICT    312    312       W -> R (IN REF. 2).
SQ   SEQUENCE   482 AA;  55103 MW;  4D35B7C1ED7838D6 CRC64;


  Query Match             56.7%;  Score 1544;  DB 1;  Length 482;
  Best Local Similarity   59.4%;  Pred. No. 2.6e-112;
  Matches  291;  Conservative   67;  Mismatches   98;  Indels   34;  Gaps    7;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | :||||||:|| :|||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 GTRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::|||||||||||||: |||   |||| :||||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVKLN 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
           |||:|||||||||| |||||||||||||||||| |:|||||:|||||||||||
Db     126 KQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGVEE 185
```

```
Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| |||||||||||:|:|||||| ::||| | ||||::| || |||||||| :|||:|
Db     186 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVM 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           ||||:|:| |||||||:|||||||||||||:|||||:||:|::|||:|:|:|||||||
Db     246 SKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGYTI 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           ||||||| ||| ||| |: :::| ::|:|||||: ||::|||| |:|: : ||:|:
Db     306 RNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIKQR 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD--------------PDSDM 418
           |  ||  | ||| |  |  | |    |  |||::| |||                 ||:
Db     366 LFENLRMLPHAPGVQMQAIPEDAIPEESGDEDEDDPDKRISICSSDKRIACEEEFSDSEE 425

Qy     419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
           : :  ||   :     ||| |    | |:||      |  |    || |   ||
Db     426 EGEGGRKNSSNFKKAKRVKTE----DEKEKD-----PEEKK----EVTEEEKTKEEKPEA 472

Qy     474 VGVEEASVKM 483
           ||:|  ||:
Db     473 KGVKE-EVKL 481


RESULT   5
HDA1_CHICK
ID   HDA1_CHICK     STANDARD;      PRT;   480 AA.
AC   P56517;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1 OR HDAC1A.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takami Y.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sun J.M., Chen H.Y., Davie J.R.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: FORMS A COMPLEX WITH RBAP48 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF039751; AAB96923.1; -.
DR   EMBL; AF043328; AAB99850.1; -.
DR   EMBL; AF044169; AAC00504.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      299    302       POLY-GLY.
FT   DOMAIN      396    399       POLY-GLU.
FT   DOMAIN      446    453       POLY-GLU.
FT   CONFLICT      3      4       LT -> VM (IN REF. 2; AAB99850).
SQ   SEQUENCE   480 AA;  54938 MW;  82C78CE285C779D9 CRC64;


  Query Match           56.5%;  Score 1538;  DB 1;  Length 480;
  Best Local Similarity   57.9%;  Pred. No. 7.4e-112;
  Matches  282;  Conservative   74;  Mismatches  107;  Indels   24;  Gaps    4;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | |||||:|| :|||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 GTKRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: |||||  |:   :  :|::|||||||||||||: |||  |||||  :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSAGGSVASAVKLN 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
               |||:|||||||||| |||||||||||||||||| |:||||:|||||||||||
Db     126 KQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGVEE 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| |||||||||||:|:|||||| ::||| | ||||::| || |||||||| :|||::
Db     186 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVI 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           |||| |:| |||||||:||||||||||||:|||||:||:|::|||:|:|:|||||||
Db     246 SKVMETFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGYTI 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           ||||||| ||| |||  |: :::| ::|:|||||: ||::|||| |:|: : ||:|:
Db     306 RNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIKQR 365

Qy     374 LLHNLSKLQHAPSVPFQERPPD-TETPEVDEDQEDGDKRWD-PDSDMDVDDDRKPIPSRV 431
           |  ||  | ||| |  |  | |  :    ||::|| :||    :|| :  |
Db     366 LFENLRMLPHAPGVQMQPIPEDAVQEDSGDEEEEDPEKRISIRNSDKRISCD-------- 417

Qy     432 KREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEASVKMEEEGTNKG 491
           |  : : : : | | :    |   |: :|               :|   | :|:   |
Db     418 -EEFSDSEDEGEGGRKNVANFKKAKRVKTEEE-----------KEEEEKKDEKEEEKA 463

Qy     492 GAEQAFP 498
           |:| |
Db     464 KEEKAEP 470


RESULT   6
HD12_XENLA
ID   HD12_XENLA     STANDARD;      PRT;   480 AA.
AC   O42227;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE 1-2 (HD1) (RPD3 HOMOLOG).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Patterton D., Wolffe A.P.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
```

```
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF020658; AAC60346.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      299    302       POLY-GLY.
SQ   SEQUENCE   480 AA;  54892 MW;  CA92DE34D36E39E8 CRC64;


  Query Match             55.9%;  Score 1521;  DB 1;  Length 480;
  Best Local Similarity   59.1%;  Pred. No. 1.6e-110;
  Matches  275;  Conservative   77;  Mismatches   95;  Indels   18;  Gaps    4;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | |:||||:|| :|||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 GTKKKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKASAEEMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::|||||||||||||||: |||   |||| :||||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVKLN 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
             ||::||:|||||||| ||||||||||||||||||||| |:||:|:|||||||||||
Db     126 KQQTDISVNWSGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVVYIDIDIHHGDGVEE 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| |||||:||||||:|:|||||| ::||| | ||||::| || ||||||||  :|||:|
Db     186 AFYTTDRVMSVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVM 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           |||||:|:| |||||||||||||||||||||:|||||:||:|:::::||:|:|:||||||||
Db     246 TKVMEMFQPSAVVLQCGADSLSGDRLGCFNLTIKGHAKCVEFIKTFNLPMLMLGGGGYTI 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           ||||||| ||| |||  |: :::| ::|:|||||: ||::|||| |:|: : ||:|:
Db     306 RNVARCWTYETAVALDSEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIKQR 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDT-ETPEVDEDQEDGDKRWD----------PDSDM 418
           |  ||  | ||| |  |  | |:      :|||:|| |||                ||:
Db     366 LFENLRMLPHAPGVQMQAIPEDSVHDDSGEEDEEDPDKRISIRSSDKRIACDEEFSDSED 425

Qy     419 DVDDDRKPIPS--RVKREAVEPDTKDKDGLKGIMERGKGCEVEVD 461
           : :  || : :  :|||   | : |: :  | : |  |  : : |
Db     426 EGEGGRKNVANFKKVKRVKTE-EEKEGEDKKDVKEEEKAKDEKTD 469


RESULT   7
HD11_XENLA
ID   HD11_XENLA     STANDARD;      PRT;   480 AA.
AC   Q91695;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE 1-1 (HD1) (MATERNALLY-EXPRESSED
DE   HISTONE DEACETYLASE) (HDM) (AB21).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocyte;
RX   MEDLINE=98036059; PubMed=9370292;
RA   Ladomery M.R., Lyons S., Sommerville J.;
RT   "Xenopus HDm, a maternally expressed histone deacetylase, belongs to
RT   an ancient family of acetyl-metabolizing enzymes.";
RL   Gene 198:275-280(1997).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: OOCYTE.
CC   -!- DEVELOPMENTAL STAGE: ACCUMULATES IN PREVITELLOGENIC OOCYTES AND IS
CC       MAINTAINED AT CONSTANT LEVEL THROUGHOUT OOGENESIS AND INTO EARLY
CC       EMBRYOGENESIS. DECLINES THROUGH GASTRULA TO NEURULA. NOT
CC       DETECTABLE BETWEEN NEURULA AND TAILBUD, NOR IN ADULT TISSUES OTHER
CC       THAN OVARY.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X78454; CAA55211.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      299    302       POLY-GLY.
SQ   SEQUENCE   480 AA;  54747 MW;  7B831822235DADB5 CRC64;


  Query Match             55.8%;  Score 1520;  DB 1;  Length 480;
  Best Local Similarity   59.8%;  Pred. No. 1.9e-110;
  Matches  278;  Conservative   73;  Mismatches   96;  Indels   18;  Gaps    5;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | |:||||:|| :|||||||||||||||||||| || :||| : |:: :|  |   |:
Db       6 GTKKKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIFRPHKASAEDMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::|||||||||||||||: |||  |||||  :||||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSAGGSVASAVKLN 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
             ||::||:|||||||| ||||||||||||||||||||| |:||:|:|||||||||||
Db     126 KQQTDISVNWSGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVVYIDIDIHHGDGVEE 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| |||||:||||||:|:|||||| ::||| | ||||::|  | |||||||  :|||:|
Db     186 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYALRDGIDDESYEAIFKPVM 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           ||||:|:| |||||||||||||||||||||:|||||:||:|:::::||:|:|:||||||||
Db     246 SKVMEMFQPSAVVLQCGADSLSGDRLGCFNLTIKGHAKCVEFIKTFNLPLLMLGGGGYTI 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           ||||||| ||| |||  |: :::| ::|:|||||: ||::|||| |:|: : ||:|:
Db     306 RNVARCWTYETAVALDSEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIKQR 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDT---ETPEVDEDQED-------GDKRWDPDSDMDVDDD 423
           |  ||  | ||| |  |    |:   :: | |||  |         |||   | :    :|
Db     366 LFENLRMLPHAPGVQMQAVAEDSIHDDSGEEDEDDPDKRISIRSSDKRIACDEEFSDSED 425

Qy     424 -----RKPIPS--RVKREAVEPDTKDKDGLKGIMERGKGCEVEVD 461
                || : :  :|||   | : |: :  | : |  |  : : |
```

```
Db    426 EGEGGRKNVANFKKVKRVKTE-EEKEGEDKKDVKEEEKAKDEKTD 469


RESULT   8
HDA2_MOUSE
ID   HDA2_MOUSE     STANDARD;      PRT;   488 AA.
AC   P70288;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 2 (HD2) (YY1 TRANSCRIPTION FACTOR BINDING
DE   PROTEIN).
GN   HDAC2 OR YY1BP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymphoma;
RX   MEDLINE=97075080; PubMed=8917507;
RA   Yang W.-M., Inouye C.J., Zeng Y., Bearss D., Seto E.;
RT   "Transcriptional repression by YY1 is mediated by interaction with a
RT   mammalian homolog of the yeast global regulator RPD3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
CC       AT LEAST WITH THE ZINC-FINGER TRANSCRIPTION FACTOR YY1.
CC   -!- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX WITH YY1.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31758; AAC52889.1; -.
DR   MGD; MGI:1097691; Hdac2.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      300    303       POLY-GLY.
SQ   SEQUENCE   488 AA;  55302 MW;  B9843D24A775157C CRC64;


  Query Match           55.6%;  Score 1513.5;  DB 1;  Length 488;
  Best Local Similarity   55.7%;  Pred. No. 6.1e-110;
  Matches  280;  Conservative   82;  Mismatches  106;  Indels   35;  Gaps    7;

Qy       8 LASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPA 67
           :|    | |:||||:|| ::|||||||||||||||||||| || :||| : |:: :|  |
Db       1 MAYSQGGGKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKA 60

Qy      68 RERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVG 127
             :: ::|:|:|: ||||| |:   :  :|::|||||||||||||: |||   ||||
Db      61 TAEEMTKYHSDEYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVA 120

Qy     128 GSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHH 187
           |:||||    |:|:|||||||||| ||||||||||||||||||| |:||||:||||||
Db     121 GAVKLNRQQTDMAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHH 180

Qy     188 GDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHL 247
           ||||||||| |||||||||||:|:|||||| ::||| | ||||::| |: ||||||||
Db     181 GDGVEEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNFPMRDGIDDESYGQ 240

Qy     248 LFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLG 307
           :|||||: ||||:::| |||||||||||||||||||||::||||:||:  ::|||:|||:||
Db     241 IFKPIISKVMEMYQPSAVVLQCGADSLSGDRLGCFNLTVKGHAKCVEVAKTFNLPLLMLG 300

Qy     308 GGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQML 367
           |||||||||||||| ||| |||  |: :::| ::|:|||||||: ||::|||| |:|: : :
Db     301 GGGYTIRNVARCWTYETAVALDCEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTPEYM 360

Qy     368 EEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDED--QEDGDKRWDPDSDMDVDDDRK 425
           |:|:  |  ||  | ||| |  |  | |     | ||   |||:   |||        |
Db     361 EKIKQRLFENLRMLPHAPGVQMQAIPEDA----VHEDSGDEDGE---DPDK-------RI 406

Qy     426 PIPSRVKREAVEPDTKDKD-----GLKGIMERGKGCE---VEVDESGSTKVTGVNPVGVE 477
             | :  || | : :  | :     | : : :  || :   :| |:  :            |
Db     407 SIRASDKRIACDEEFSDSEDEGEGGRRNVADHKKGAKKARIEEDKKET----------E 455

Qy     478 EASVKMEEEGTNKGGAEQAFPPK 500
           :    ::||  :|  : :   ||
Db     456 DKKTDVKEEDKSKDNSGEKTDPK 478


RESULT   9
HDA1_STRPU
ID   HDA1_STRPU     STANDARD;      PRT;   576 AA.
AC   P56518;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nemer M.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF032919; AAB87685.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   576 AA;  64078 MW;  B3D11A844A2088E9 CRC64;


  Query Match           55.4%;  Score 1507;  DB 1;  Length 576;
  Best Local Similarity   56.0%;  Pred. No. 2.4e-109;
  Matches  280;  Conservative   75;  Mismatches  119;  Indels   26;  Gaps    6;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
```

```
Db      5 GTKKRVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLILNYGLYRKMEIYRPHKAVMEEMT 64

Qy     74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133

Db     65 KYHSDDYVKFLRTIRPDNMSEYTKQMQRFNVGEDCPVFDGLYEFCQLSSGGSVAGAVKLN 124

Qy    134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193

Db    125 KQQTDIAINWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGVEE 184

Qy    194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253

Db    185 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNFPLRDGIDDESYDKIFKPIM 244

Qy    254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313

Db    245 CKVMEMYQPSAICLQCGADSLSGDRLGCFNLTLKGHAKCVEFMKQYNLPLLLMGGGGYTI 304

Qy    314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373

Db    305 RNVARCWTYETSTALGVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTGEYLDKIKTR 364

Qy    374 LLHNLSKLQHAPSVPFQERPPDTETPEVDEDQED----------GDKRWDPDSDMDV 420

Db    365 LYENMRMIPHAPGVQMQPIPEDA-IPD-DSDAEDEAENPDKRISIMAQDKRIQRDDEFSD 422

Qy    421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480

Db    423 SEDEG--ETRLPGEGGRRDHRSHKAKRSKIDDSPGKEADKEAKSSDASKEAKP-AAEPQA 479

Qy    481 VKMEEEGTNKGGAEQAFPPK 500

Db    480 VPMD--------TTPAPPPK 491
```

RESULT 10
HDA2_HUMAN
ID HDA2_HUMAN STANDARD; PRT; 488 AA.
AC Q92769;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE 2 (HD2).
GN HDAC2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=97075080; PubMed=8917507;
RA Yang W.-M., Inouye C.J., Zeng Y., Bearss D., Seto E.;
RT "Transcriptional repression by YY1 is mediated by interaction with a
RT mammalian homolog of the yeast global regulator RPD3.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
CC -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -!- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
CC WITH MAD, SIN3, YY1 AND N-COR.
CC -!- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX AT LEAST WITH YY1.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; LOWER LEVEL IN BRAIN AND
CC LUNG.
CC -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC ----------
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ----------
DR EMBL; U31814; AAC50814.1; -.
DR MIM; 605164; -.
DR InterPro; IPR000286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
DR PRINTS; PR01271; HISDACETLASE.
KW Hydrolase; Nuclear protein.
FT DOMAIN 300 303 POLY-GLY.
SQ SEQUENCE 488 AA; 55325 MW; 3AE554ADC7734B70 CRC64;

Query Match 54.9%; Score 1494; DB 1; Length 488;
Best Local Similarity 54.3%; Pred. No. 2e-108;
Matches 277; Conservative 80; Mismatches 113; Indels 40; Gaps 6;

```
Qy      8 LASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPA 67

Db      1 MAYSQGGGKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKA 60

Qy     68 RERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVG 127

Db     61 TAEEMTKYHSDEYIKFLRSIRPDNMSEYSKQMHIFNVGEDCPAFDGLFEFCQLSTGGSVA 120

Qy    128 GSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHH 187

Db    121 GAVKLNRQQTDMAVNWAGGLHHAKKYEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHH 180

Qy    188 GDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHL 247

Db    181 GDGVEEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNFPMCDGIDDESYGQ 240

Qy    248 LFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLG 307

Db    241 IFKPIISKVMEMYQPSAVVLQCGADSLSGDRLGCFNLTVKGHAKCVEVVKTFNLPLLMLG 300

Qy    308 GGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQML 367

Db    301 GGGYTIRNVARCWTYETAVALDCEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTPEYM 360

Qy    368 EEIRNDLLHNLSKLQHAPSVPFQERPPD-TETPEVDEDQEDGDKRWD---------- 413

Db    361 EKIKQRLFENLRMLPHAPGVQMQAIPEDAVHEDSGDEDGEDPDKRISIRASDKRIACDEE 420

Qy    414 -PDSDMDVDDDRKPIPSR---VKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVT 469

Db    421 FSDSEDEGEGGRRNVADHKKGAKKARIEEDKKETEDKK----------TDVKEEDKSK-- 468

Qy    470 GVNPVGVEEASVKMEEEGTNKGGAEQAFPP 499

Db    469 -------DNSGEKTDTKGTK---SEQLSNP 488
```

RESULT 11
HDAC_DROME
ID HDAC_DROME STANDARD; PRT; 520 AA.
AC Q94517; O17429;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROBABLE HISTONE DEACETYLASE (HD) (DRPD3).
GN RPD3 OR HDAC1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

```
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=97113416; PubMed=8955276;
RA   de Rubertis F., Kadosh D., Henchoz S., Pauli D., Reuter G., Struhl K.,
RA   Spierer P.;
RT   "The histone deacetylase RPD3 counteracts genomic silencing in
RT   Drosophila and yeast.";
RL   Nature 384:589-591(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Johnson C.A., White D., O'Neill L.P., Turner B.M.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY). IN DROSOPHILA, IT IS INVOLVED IN POSITION-EFFECT
CC       VARIEGATION (PEV) AND IS DIRECTLY INVOLVED IN COUNTERACTING
CC       GENOME SILENCING.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y09258; CAA70455.1; -.
DR   EMBL; AF026949; AAC23917.1; -.
DR   FlyBase; FBgn0015805; Rpd3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   CONFLICT     50     51       DI -> EIY (IN REF. 2).
FT   CONFLICT     66     66       C -> S (IN REF. 2).
FT   CONFLICT     96     96       D -> N (IN REF. 2).
FT   CONFLICT    105    105       E -> D (IN REF. 2).
FT   CONFLICT    295    295       V -> VV (IN REF. 2).
FT   CONFLICT    370    370       L -> V (IN REF. 2).
FT   CONFLICT    506    506       S -> T (IN REF. 2).
SQ   SEQUENCE   520 AA;  58183 MW;  D02EA3DBD3C64688 CRC64;


  Query Match             54.3%;  Score 1478;  DB 1;  Length 520;
  Best Local Similarity   56.1%;  Pred. No. 3.7e-107;
  Matches  267;  Conservative   78;  Mismatches  105;  Indels   26;  Gaps    5;

Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           |::|||:|| ::|||||||||||||||||||| || :||| : | : :|  |   :: :|
Db       6 KKRVCYYYDSDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMDI-RPHKATADEMTKF 64

Qy      76 HADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHG 135
           | |:|| ||||| |:   :  :|::|||||||||||||||| |||  |||||  :|||| 
Db      65 HCDEYVRFLRSIRPDNMSEYNKQMQRFNVGEDCPVFDGLYEFCQLSAGGSVAAAVKLNKQ 124

Qy     136 LCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAF 195
             :| ||| |||||||| |||||||||||| |||||| |:||||:|||:|||||||||| 
Db     125 ASEICINWGGGLHHAKKSEASGFCYVNDIVLGILELLKYHQRVLYIDIDVHHGDGVEEAF 184

Qy     196 YATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           | |||||||||||:|:|||||| ::||| | ||||::|:|| ||:||::|  :| ||: | 
Db     185 YTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNIPLRDGMDDDAYESIFVPIISK 244

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           ||| |:| |||||||||||:|||||||||::||| :||:|:: :|:| |::|||||||| 
Db     245 VMETFQPAAVVLQCGADSLTGDRLGCFNLTVKGHGKCVEFVKKYNLPFLMVGGGGYTIRK 304
```

```
Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLL 375
           |:||| ||| ||| ||: :::| ::|:|||||: ||::|||| |:|: : ||:|:| |
Db     305 VSRCWTYETSVALAVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTSEYLEKIKNRLF 364

Qy     376 HNLSKLQHAPSVPFQERPPDTETPEVDED----------QEDGDKRWDPDSD-MDVDD-- 422
           || | ||| |  |  | |    | |::          | | |||  |:::  | :| 
Db     365 ENLRMLPHAPGVQIQAIPEDAINDESDDEDKVDKDDRLPQSDKDKRIVPENEYSDSEDEG 424

Qy     423 -----DRKPIPSRVKREAVEPDTKDKDG-------LKGIMERGKGCEVEVDESGST 466
                | :    : ||  :: ||            :|   |:| | : |    : :| 
Db     425 EGGRRDNRSYKGQRKRPRLDKDTNSNKASSETSSEIKDEKEKGDGADGEESTASNT 480


RESULT  12
HDA2_CHICK
ID   HDA2_CHICK     STANDARD;      PRT;   488 AA.
AC   P56519;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 2 (HD2).
GN   HDAC2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takami Y.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039752; AAB96924.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      300    303       POLY-GLY.
SQ   SEQUENCE   488 AA;  55153 MW;  4F79B9C0D4A2D065 CRC64;


  Query Match             53.8%;  Score 1465.5;  DB 1;  Length 488;
  Best Local Similarity   55.4%;  Pred. No. 3.2e-106;
  Matches  274;  Conservative   81;  Mismatches  107;  Indels   33;  Gaps    7;

Qy       8 LASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPA 67
           :|    | |:||||:|| ::|||||||||||||||||||| || :||| : |:: :|  |
Db       1 MAYSQGGGKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKA 60

Qy      68 RERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVG 127
             :: ::|:|:|: ||||| |:   :  :|::|||||||||||||: |||   ||||
Db      61 TAEEMTKYHSDEYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVA 120

Qy     128 GSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHH 187
           |:||||    |:|:|||||||||| |||||||||||||||||| |:||||:||||||
Db     121 GAVKLNRQQTDMAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHH 180
```

```
Qy    188 GDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHL 247
          ||||||||| ||||||||     : ||||| ::||| | ||||::| |: ||||||||
Db    181 GDGVEEAFYTTDRVMTVSEVSMVNNFPGTGDLRDIGAGKGKYYAVNFPMRDGIDDESYGQ 240

Qy    248 LFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLG 307
          :|||||: ||||:::| ||||||||||||||||||||||::||||:||: :::|||:|||:||
Db    241 IFKPIISKVMEMYQPSAVVLQCGADSLSGDRLGCFNLTVKGHAKCVEVVKTFNLPLLMLG 300

Qy    308 GGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQML 367
          |||||||||||||| ||| ||||  |: :::| ::|:|||||: ||::|||| |:|: : :
Db    301 GGGYTIRNVARCWTYETAVALDCEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTPEYM 360

Qy    368 EEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDED--QEDGDKRWDPDSDMDVDDDRK 425
          |:|:  |  ||  | ||| |  |  | |     | ||   |||:   |||         |
Db    361 EKIKQRLFENLRMLPHAPGVQMQAIPEDA----VHEDSGDEDGE---DPDK-------RI 406

Qy    426 PIPSRVKREAVEPDTKDKD-----GLKGIMERGKGCE---VEVDESGSTKVTGVNPVGVE 477
            | :  || | : :  | :     | : : :  || :   :| |:  :            :
Db    407 SIRASDKRIACDEEFSDSEDEGEGGRRNVADHKKGAKKARIEEDKKETED---------K 457

Qy    478 EASVKMEEEGTNKGG 492
          :| || |::  :  |
Db    458 KADVKEEDKSKDNSG 472


RESULT  13
HDA1_CAEEL
ID   HDA1_CAEEL     STANDARD;      PRT;   461 AA.
AC   O17695;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE 1.
GN   HDA-1 OR C53A5.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Mortimore B.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z81486; CAB03984.1; -.
DR   WormPep; C53A5.3; CE08952.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   461 AA;  52137 MW;  9A104E88C5A1C07A CRC64;


  Query Match           51.1%;  Score 1392;  DB 1;  Length 461;
  Best Local Similarity  57.2%;  Pred. No. 1.5e-100;
  Matches  259;  Conservative   68;  Mismatches  102;  Indels   24;  Gaps    4;

Qy      1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
            |:: |  :  |    ||:| |:||  :|||||||||| |||||||||| |: :|:: ::::
Db      1 MNSNGPLMEHG----KRRVAYYYDSNIGNYYYGQGHVMKPHRIRMTHHLVLNYGLYRNLE 56

Qy     61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
          : :||||   |: |||:|:|::||:|  |:  :   :|: :|||||||:|||| |||
Db     57 IFRPFPASFEDMTRFHSDEYMTFLKSANPDNLKSFNKQMLKFNVGEDCPLFDGLYEFCQL 116

Qy    121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           :|||:  : |||     |||||| |||||||| ||||||| ||||| |||||| |:||||
Db    117 SSGGSLAAATKLNKQKVDIAINWMGGLHHAKKSEASGFCYTNDIVLGILELLKYHKRVLY 176

Qy    181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
          ||||:|||||||||| ||||||||||:||:||||| ::||| | || ||:|||| |||
Db    177 VDIDVHHGDGVEEAFYTTDRVMTVSFHKYGDFFPGTGDLKDIGAGKGKLYSVNVPLRDGI 236

Qy    241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
          | ||  :||||| |||| | | ||||||||||:||||| |||::||| || :| ||:|
Db    237 TDVSYQSIFKPIMTKVMERFDPCAVVLQCGADSLNGDRLGPFNLTLKGHGECARFFRSYN 296

Qy    301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
          |||:::|||||| ||||||| ||| :|:  || :::| ::|:|||||:| ||:  ||  |
Db    297 VPLMMVGGGGYTPRNVARCWTYETSIAVDKEVPNELPYNDYFEYFGPNYRLHIESSNAAN 356

Qy    361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDE----DQEDGDKRWDP-- 414
          :||  || ::: |:: || :|   |||  :  | |  :   |:    || : |||  |
Db    357 ENSSDMLAKLQTDVIANLEQLTFVPSVQMRPIPEDALSALNDDSLIADQANPDKRLPPQI 416

Qy    415 ------------DSDMDVDDDRKPIPSRVKREA 435
                       | : || |    |  || |
Db    417 TDGMIQDDGDFYDGEREGDDRRN--ESDAKRAA 447


RESULT  14
HDA3_CHICK
ID   HDA3_CHICK     STANDARD;      PRT;   428 AA.
AC   P56520;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 3 (HD3).
GN   HDAC3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takami Y.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039753; AAB96925.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
```

```
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   428 AA;  48901 MW;  42E32733AD2BBF07 CRC64;


  Query Match             50.6%;  Score 1376;  DB 1;  Length 428;
  Best Local Similarity   56.8%;  Pred. No. 2.3e-99;
  Matches  242;  Conservative   74;  Mismatches   88;  Indels   22;  Gaps    4;

Qy      17 RKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFH 76
           : | |||||:|||::|| |||||||||: :|||:|: |||| : | | ||: | : |:|||||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      77 ADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
           ::|||: ||: ::|   |   : |  ||||:||||| ||: ||  | | |: |: :|||: :
Db      63 SEDYIDFLQRVSPNNMQGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNNKI 122

Qy     137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
           |||||||||||||||| ||||||||||||: |||||| | ||||:||||||||||:||||
Db     123 CDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEAFY 182

Qy     197 ATDRVMTVSFHKFGDY-FPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           ||||||||||||:|:| ||||| : ::|  ||:||:||||| |||||:||  ||:|:: :
Db     183 LTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYALNVPLRDGIDDQSYKHLFQPVINQ 242

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           |:: ::|  :|||||||||  ||||||||||:|| |||:::::|||:|||:||||||:||
Db     243 VVDYYQPTCIVLQCGADSLGRDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN 302

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRNDL 374
           ||||| ||| : :   : :::|  ||:||| ||:|||    : :||:|||| |::||  :
Db     303 VARCWTYETSLLVDEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQTI 362

Qy     375 LHNLSKLQHAPSVPFQERP----------PDTE---------TPEVDEDQEDGDKRWDP 414
             ||  | |||||   : |           || |          ||   :  |||   |
Db     363 FENLKMLNHAPSVQIHDVPSDLLSYDRTDEPDPEERGSEENYSRPEAANEFYDGDHDNDK 422

Qy     415 DSDMDV 420
           :||:::
Db     423 ESDVEI 428


RESULT  15
HDA3_HUMAN
ID   HDA3_HUMAN     STANDARD;      PRT;   428 AA.
AC   O15379; O43268; Q9UEV0; Q9UEI5;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 3 (HD3) (RPD3-2).
GN   HDAC3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=T-cell, and Spleen;
RX   MEDLINE=98125547; PubMed=9464271;
RA   Dangond F., Hafler D.A., Tong J.K., Randall J., Kojima R., Utku N.,
RA   Gullans S.R.;
RT   "Differential display cloning of a novel human histone deacetylase
RT   (HDAC3) cDNA from PHA-activated immune cells.";
RL   Biochem. Biophys. Res. Commun. 242:648-652(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Fibroblast;
RX   MEDLINE=98010646; PubMed=9346952;
RA   Yang W.-M., Yao Y.-L., Sun J.-M., Davie J.R., Seto E.;
RT   "Isolation and characterization of cDNAs corresponding to an
RT   additional member of the human histone deacetylase gene family.";
RL   J. Biol. Chem. 272:28001-28007(1997).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=98169450; PubMed=9501169;
RA   Emiliani S., Fischle W., van Lint C., Al-Abed Y., Verdin E.;
RT   "Characterization of a human RPD3 ortholog, HDAC3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:2795-2800(1998).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=99162406; PubMed=10051405;
RA   Mahlknecht U., Emiliani S., Najfeld V., Young S., Verdin E.;
RT   "Genomic organization and chromosomal localization of the human
RT   histone deacetylase 3 gene.";
RL   Genomics 56:197-202(1999).
RN   [5]
RP   SEQUENCE OF 95-353 AND 407-428 FROM N.A.
RA   Lynch E.D., Lee M.K., King M.-C.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS.
CC   -!- FUNCTION: MAY PARTICIPATE TO THE REGULATION OF TRANSCRIPTION
CC       THROUGH ITS BINDING WITH THE ZINC-FINGER TRANSCRIPTION FACTOR YY1;
CC       INCREASES YY1 REPRESSION ACTIVITY.
CC   -!- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX AT LEAST WITH YY1.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1/RPD3-2B (SHOWN HERE) AND
CC       2/RPD3-2A; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U66914; AAC52038.1; -.
DR   EMBL; U75697; AAB88241.1; -.
DR   EMBL; U75696; AAB88240.1; -.
DR   EMBL; AF005482; AAB87752.1; -.
DR   EMBL; AF039703; AAC98927.1; -.
DR   EMBL; AF059650; AAC26509.1; -.
DR   EMBL; AF053138; AAC08351.1; -.
DR   EMBL; AF053137; AAC08351.1; JOINED.
DR   EMBL; AF053139; AAC08352.1; -.
DR   MIM; 605166; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein; Alternative splicing.
FT   VARSPLIC      1     15       MAKTVAYFYDPDVGN -> MIVFKPYQASQHDMCR (IN
FT                                ISOFORM 2).
FT   CONFLICT    359    359       R -> L (IN REF. 1).
SQ   SEQUENCE   428 AA;  48847 MW;  94485C1EBDCF5AD0 CRC64;


  Query Match             50.4%;  Score 1373;  DB 1;  Length 428;
  Best Local Similarity   56.6%;  Pred. No. 4e-99;
  Matches  241;  Conservative   73;  Mismatches   90;  Indels   22;  Gaps    3;

Qy      17 RKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFH 76
           : | |||||:|||::|| |||||||||: :|||:|: |||| : | | ||: | : |:|||||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      77 ADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
           ::|||: ||: ::|   |   : |  ||||:||||| ||: ||  | | |: |: :|||: :
Db      63 SEDYIDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNNKI 122
```

```
Qy     137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
           ||||||||||||||||| ||||||||||||: |||||| | ||||:|||||||||||:|||| 
Db     123 CDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEAFY 182

Qy     197 ATDRVMTVSFHKFGDY-FPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           ||||||||||:|:| ||||| : ::|  ||:|| ||||| |||||:||  ||:|:: : 
Db     183 LTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVINQ 242

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           |:: ::|  :|||||||||  ||||||||||:|| |||::::|||:|||:|||||||:|| 
Db     243 VVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN 302

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRNDL 374
           ||||| ||| : :   : :::|  ||:||| ||:|||    : :|||:|||| |::||  : 
Db     303 VARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQTI 362

Qy     375 LHNLSKLQHAPSVPFQERPPDTET--------------------PEVDEDQEDGDKRWDP 414
             ||  | |||||   : | |  |                     ||   :  |||   | 
Db     363 FENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDNDK 422

Qy     415 DSDMDV 420
           :||:::
Db     423 ESDVEI 428
```

Search completed: April 26, 2002, 17:57:50
Job time: 333 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:50:57 ; Search time 25.88 Seconds
(without alignments)
1474.631 Million cell updates/sec

Title: US-09-645-337-2
Perfect score: 2722
Sequence: 1 MDTGGNSLASGPDGVKRKVC.....KMEEEGTNKGGAEQAFPPKT 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2719	99.9	501	2	T05640	histone deacetylas
2	1986.5	73.0	513	2	T01413	probable histone d
3	1520	55.8	480	1	S60381	RPD3 protein homol
4	1392	51.1	461	2	T20163	hypothetical prote
5	1366	50.2	424	2	JC7102	histone deacetylas
6	1366	50.2	428	2	JC5834	histone deacetylas
7	1341	49.3	433	1	S22284	transcription regu
8	1336	49.1	405	2	T40300	histone deacetylas
9	1334	49.0	465	2	T23963	hypothetical prote
10	1298	47.7	419	2	T47443	hypothetical prote
11	1167.5	42.9	434	2	T11643	histone deacetylas
12	1077	39.6	507	2	T19067	hypothetical prote
13	1066.5	39.2	452	2	S64211	probable transcrip
14	475	17.5	375	2	D70388	acetoin utilizatio
15	471.5	17.3	470	2	S54089	hypothetical prote
16	448.5	16.5	389	2	E84054	acetoin dehydrogen
17	443.5	16.3	387	1	S39643	acetate / acetoin
18	425.5	15.6	367	2	T36278	hypothetical prote
19	367.5	13.5	158	2	T48929	hypothetical prote
20	352	12.9	359	1	B69266	acetylpolyamine am
21	320.5	11.8	706	2	S62933	hypothetical prote
22	317.5	11.7	380	2	D83174	probable acetylpol
23	301.5	11.1	425	2	G86217	protein T27G7.14 [
24	299.5	11.0	142	2	T47441	hypothetical prote
25	297	10.9	310	1	A70481	acetoin utilizatio
26	296.5	10.9	1063	2	T46284	hypothetical prote
27	278	10.2	687	2	T43797	probable histone d
28	276.5	10.2	334	2	B75095	probable histone d
29	276	10.1	878	2	T17245	hypothetical prote
30	275.5	10.1	1095	2	T13964	probable histone d
31	271.5	10.0	338	2	H84173	acetoin utilizatio
32	264	9.7	335	1	H71071	hypothetical prote
33	261.5	9.6	310	2	B47050	glnA 3'-region hyp
34	243.5	8.9	796	2	T32425	hypothetical prote
35	238	8.7	331	1	C69026	acetylpolyamine am
36	236	8.7	517	2	T27101	hypothetical prote
37	230.5	8.5	385	2	A81926	hypothetical prote
38	230	8.4	343	1	G64366	acetylpolyamine am
39	227	8.3	306	2	D82126	histone deacetylas
40	224.5	8.2	369	2	F81178	histone deacetylas
41	223.5	8.2	359	2	T05998	hypothetical prote
42	208	7.6	304	1	S74557	acetylpolyamine am
43	201	7.4	301	2	H75470	histone deacetylas
44	197	7.2	344	2	B83605	probable acetylpol
45	196.5	7.2	346	2	H83469	acetylpolyamine am

ALIGNMENTS

RESULT 1
T05640
histone deacetylase (EC 3.5.1.-) F20D10.250 - Arabidopsis thaliana
N;Alternate names: protein F20D10.250
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 12-Nov-1999
C;Accession: T05640
R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15420
A;Accession: T05640
A;Molecule type: DNA
A;Residues: 1-501 <BEV>
A;Cross-references: EMBL:AL035538; GSPDB:GN00062; ATSP:F20D10.250
A;Experimental source: cultivar Columbia; BAC clone F20D10
C;Genetics:
A;Gene: ATSP:F20D10.250
A;Map position: 4
A;Introns: 176/3; 323/3; 400/3; 424/2; 443/3; 467/3
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase
F;21-320/Domain: RPD3/acuC homology <RAH>

Query Match 99.9%; Score 2719; DB 2; Length 501;
Best Local Similarity 99.8%; Pred. No. 1.1e-202;
Matches 500; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
          |||||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 VLKPFPARDRDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120

Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
```

```
Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420

Qy     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480

Qy     481 VKMEEEGTNKGGAEQAFPPKT 501
           |||||||||||||||||||||
Db     481 VKMEEEGTNKGGAEQAFPPKT 501


RESULT   2
T01413
probable histone deacetylase (EC 3.5.1.-) rpd3 - maize
C;Species: Zea mays (maize)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 12-Nov-1999
C;Accession: T01413
R;Rossi, V.; Hartings, H.; Motto, M.
Mol. Gen. Genet. 258, 288-296, 1998
A;Title: Identification and characterisation of an RPD3 homologue from maize (Zea mays L
A;Reference number: Z14321; MUID:98307342
A;Accession: T01413
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-513 <ROS>
A;Cross-references: EMBL:AF035815; NID:g2665839; PIDN:AAC50038.1; PID:g2665840
A;Experimental source: strain W22
C;Genetics:
A;Gene: rpd3
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase
F;27-325/Domain: RPD3/acuC homology <RAH>


  Query Match             73.0%;  Score 1986.5;  DB 2;  Length 513;
  Best Local Similarity   75.2%;  Pred. No. 5.7e-146;
  Matches  377;  Conservative   44;  Mismatches   61;  Indels   19;  Gaps    7;

Qy       3 TGGNSLAS-GPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
           :|||||  | |||| ||:|||||||:|||||||||||||||||||||:|||  ||||   |||
Db       8 SGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPMKPHRIRMTHSLLARYGLLNQMQV 67

Qy      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTY 121
           :| |||||:|||||||::|:::||||:|||||||||| ||||||||:||| ||||||||||
Db      68 YRPNPARERELCRFHAEEYINFLRSVTPETQQDQIRLLKRFNVGEECPVLDGLYSFCQTY 127

Qy     122 AGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYV 181
           || ||||:|| |||  ||||||:|||||||||||||||||||||||||| |||||||
Db     128 AGASVGGAVKFNHG-HDIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKHHERVLYV 186

Qy     182 DIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGID 241
           |||||||||||||| ||||||||||||||||||| |:|||:  ||||||||||||||||
Db     187 DIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGHSKGKYYSLNVPLDDGID 246

Qy     242 DESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNV 301
           ||||  ||||||||||:|||||||||||||||||||||||||||||||||::||||||
Db     247 DESYQSLFKPIMGKVMEVFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYMRSFNV 306

Qy     302 PLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENK 361
           |||||||||||||||||||||||||||| | ||||| :||||||||||||||||||||
Db     307 PLLLLGGGGYTIRNVARCWCYETGVALGQEPEDKMPVNEYYEYFGPDYTLHVAPSNMENK 366

Qy     362 NSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVD 421
           |:|| |::||:     ||||:||||| |||| |||| || ||||:| |:| ||||||:||
Db     367 NTRQQLDDIRS----KLSKLRHAPSVHFQERVPDTEIPEQDEDQDDPDERHDPDSDMEVD 422

Qy     422 DD--------RKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK---VTG 470
           |         |  :  ::|||  |  |: :|| : :    :| |   :: ||:|
Db     423 DHKAVEESSRRSILGIKIKREFGENATRVQDGGR-VASEHRGLEPMAEDIGSSKQAPQAD 481

Qy     471 VNPVGVEEAS-VKMEEEGTNK 490
           .
             .
```

```
           : : ::| | || | | : |
Db     482 ASAMAIDEPSNVKNEPESSTK 502


RESULT   3
S60381
RPD3 protein homolog - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S60381; S43160
R;Ladomery, M.R.; Lyons, S.; Sommerville, J.
submitted to the EMBL Data Library, December 1994
A;Reference number: S60381
A;Accession: S60381
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-480 <LAD>
A;Cross-references: EMBL:X78454; NID:g773397; PIDN:CAA55211.1; PID:g602098
C;Superfamily: RPD3 protein; RPD3/acuC homology
F;13-312/Domain: RPD3/acuC homology <RAH1>


  Query Match             55.8%;  Score 1520;  DB 1;  Length 480;
  Best Local Similarity   59.8%;  Pred. No. 6.8e-110;
  Matches  278;  Conservative   73;  Mismatches   96;  Indels   18;  Gaps    5;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | |:||||:|| :|||||||||||||||||||| || :||| : |:: :|  |   |:
Db       6 GTKKKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIFRPHKASAEDMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::|||||||||||||: ||| ||||| :||||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSAGGSVASAVKLN 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
           ||::||:|||||||| ||||||||||||||||||| |:||:|:||||||||||
Db     126 KQQTDISVNWSGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVVYIDIDIHHGDGVEE 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| ||||||||||:|:|||||| ::||| | ||||::|  | ||||||||  :|||:|
Db     186 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYALRDGIDDESYEAIFKPVM 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           ||||:|:| ||||||||||||||||||||:|||||:||:|:::||:|||:||||||||
Db     246 SKVMEMFQPSAVVLQCGADSLSGDRLGCFNLTIKGHAKCVEFIKTFNLPLLMLGGGGYTI 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEE1RND 373
           ||||||| ||| |||  |: :::| ::|:||||||: ||::|||| |:|: : ||:|:
Db     306 RNVARCWTYETAVALDSEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEK1KQR 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDT---ETPEVDEDQED-------GDKRWDPDSDMDVDDD 423
           |  ||  | ||| |  |    |:   :: | |||  |         |||   | :    :|
Db     366 LFENLRMLPHAPGVQMQAVAEDSIHDDSGEEDEDDPDKRISIRSSDKRIACDEEFSDSED 425

Qy     424 -----RKPIPS--RVKREAVEPDTKDKDGLKGIMERGKGCEVEVD 461
                ||  :  :  :|||   | : |: :  | : |  |  : : |
Db     426 EGEGGRKNVANFKKVKRVKTE-EEKEGEDKKDVKEEEKAKDEKTD 469


RESULT   4
T20163
hypothetical protein C53A5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T20163
R;Mortimore, B.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19232
A;Accession: T20163
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-461 <WIL>
A;Cross-references: EMBL:Z81486; PIDN:CAB03984.1; GSPDB:GN00023; CESP:C53A5.3
```

```
A;Experimental source: clone C53A5
C;Genetics:
A;Gene: CESP:C53A5.3
A;Map position: 5
A;Introns: 58/3; 98/1; 387/1
C;Superfamily: RPD3 protein; RPD3/acuC homology


  Query Match            51.1%;  Score 1392;  DB 2;  Length 461;
  Best Local Similarity  57.2%;  Pred. No. 5.3e-100;
  Matches  259;  Conservative   68;  Mismatches  102;  Indels   24;  Gaps    4;

Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
           |:: |  :  |    ||:| |:||  :|||||||||| ||||||||||: |: :||| ::::
Db       1 MNSNGPLMEHG----KRRVAYYYDSNIGNYYYGQGHVMKPHRIRMTHHLVLNYGLYRNLE 56

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
           : :||||   |: |||:|:|::||:|  |:  :   :|: :|||||||||:||||| |||
Db      57 IFRPFPASFEDMTRFHSDEYMTFLKSANPDNLKSFNKQMLKFNVGEDCPLFDGLYEFCQL 116

Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           :|||:  : |||    |||||| |||||||| |||||||| ||||| |||||| |:|||||
Db     117 SSGGSLAAATKLNKQKVDIAINWMGGLHHAKKSEASGFCYTNDIVLGILELLKYHKRVLY 176

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
           ||||:|||||||||||| ||||||||||:||:|||| ::||| | || ||:|||| |||
Db     177 VDIDVHHGDGVEEAFYTTDRVMTVSFHKYGDFFPGTGDLKDIGAGKGKLYSVNVPLRDGI 236

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
           | ||  :||||| |||| | | ||||||||||:||||| |||::||| || :| ||:|
Db     237 TDVSYQSIFKPIMTKVMERFDPCAVVLQCGADSLNGDRLGPFNLTLKGHGECARFFRSYN 296

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
           |||:::|||||| |||||| ||| :|:  || :::| ::|:|||||:| ||:  ||  |
Db     297 VPLMMVGGGGYTPRNVARCWTYETSIAVDKEVPNELPYNDYFEYFGPNYRLHIESSNAAN 356

Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDE----DQEDGDKRWDP-- 414
           :||  || ::: |:: || :|   ||| :  | |  :   |:    || : |||  |
Db     357 ENSSDMLAKLQTDVIANLEQLTFVPSVQMRPIPEDALSALNDDSLIADQANPDKRLPPQI 416

Qy     415 ------------DSDMDVDDDRKPIPSRVKREA 435
                       | : : || |    |  || |
Db     417 TDGMIQDDGDFYDGEREGDDRRN--ESDAKRAA 447


RESULT   5
JC7102
histone deacetylase (EC 3.5.1.-) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C;Accession: JC7102
R;Mahlknecht, U.; Hoelzer, D.; Bucala, R.; Verdin, E.
Biochem. Biophys. Res. Commun. 263, 482-490, 1999
A;Title: Cloning and characterization of the murine histone deacetylase (HDAC3).
A;Reference number: JC7102; MUID:99423490
A;Accession: JC7102
A;Molecule type: mRNA
A;Residues: 1-424 <MAH>
A;Cross-references: GB:AF074881
C;Genetics:
A;Gene: hdac3
A;Map position: 18B3
A;Introns: 19/1; 46/3; 94/2; 121/3; 140/3; 159/2; 204/1; 231/1; 252/3; 277/2; 307/2; 327
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase; phosphoprotein
F;7-307/Domain: RPD3/acuC homology <RAH1>
F;232,282,404/Binding site: phosphate (Tyr) (covalent) #status predicted


  Query Match            50.2%;  Score 1366;  DB 2;  Length 424;
  Best Local Similarity  57.0%;  Pred. No. 4.8e-98;
  Matches  240;  Conservative   78;  Mismatches   91;  Indels   12;  Gaps    5;

Qy      17 RKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFH 76
           : | |||||:|||::|| |||||||| :|||:|: |||| : | | ||: | : |:||||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      77 ADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
           ::||: ||: ::|   |   : |  ||||:||||| ||: ||  | | |: |: :||: :
Db      63 SEDYIDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNNKI 122

Qy     137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
           |||||||||||||||| |||||||||||: |||||| | ||||:|||||||||:||||
Db     123 CDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEAFY 182

Qy     197 ATDRVMTVSFHKFGDY-FPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           |||||||||||:|:| ||||| : ::|  ||:|| ||||| |||||:||  ||:|:: :
Db     183 LTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVISQ 242

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           |:: ::|  :|||||||||  ||||||||||:|| |||:::|||:|||:|||||||:||
Db     243 VVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN 302

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRNDL 374
           ||||| ||| : :   : :::|  ||:||| ||:|||    : :|||:|||| |::||  :
Db     303 VARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQTI 362

Qy     375 LHNLSKLQHAPSVPFQERPPD----TETPEVDEDQEDGDKRWD-PDS-----DMDVDDDR 424
             ||  | ||||   : | |       | | | ::   :: :  |::     | | |:|:
Db     363 FENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDNDK 422

Qy     425 K 425
           :
Db     423 E 423


RESULT   6
JC5834
histone deacetylase (EC 3.5.1.-) 3 - human
C;Species: Homo sapiens (man)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 12-Nov-1999
C;Accession: JC5834
R;Dangond, F.; Hafler, D.A.; Tong, J.K.; Randall, J.; Kojima, R.; Utku, N.; Gullans,
Biochem. Biophys. Res. Commun. 242, 648-652, 1998
A;Title: Differential display cloning of a novel human histone deacetylase (HDAC3) cD
A;Reference number: JC5834; MUID:98125547
A;Accession: JC5834
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-428 <DAN>
A;Cross-references: GB:U66914; NID:g2934900; PIDN:AAC52038.1; PID:g2326173
C;Comment: This enzyme plays a role in multiple and complex cellular pathways of immu
gets for the treatment of cancer and autoimmunity.
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase; phosphoprotein
F;7-307/Domain: RPD3/acuC homology <RAH1>
F;232,282,404/Binding site: phosphate (Tyr) (covalent) #status predicted


  Query Match            50.2%;  Score 1366;  DB 2;  Length 428;
  Best Local Similarity  56.3%;  Pred. No. 4.9e-98;
  Matches  240;  Conservative   73;  Mismatches   91;  Indels   22;  Gaps    3;

Qy      17 RKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFH 76
           : | |||||:|||::|| |||||||| :|||:|: |||| : | | ||: | : |:||||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      77 ADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
           ::||: ||: ::|   |   : |  ||||:||||| ||: ||  | | |: |: :||: :
Db      63 SEDYIDFLQRVSPTNMQGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNNKI 122

Qy     137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
           |||||||||||||||| |||||||||||: |||||| | ||||:|||||||||:||||
Db     123 CDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEAFY 182

Qy     197 ATDRVMTVSFHKFGDY-FPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
```

```
          ||||||||||||:|:| ||||| : ::|  ||:|| ||||| |||||:||  ||:|:: :
Db    183 LTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVINQ 242

Qy    256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
          |:: ::|  :|||||||||  ||||||||||:|| |||:::::|||:|||:|||||||:||
Db    243 VVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTVRN 302

Qy    316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRNDL 374
          ||||| ||| : :   : :::|  ||:||| ||:|||    : :||:|||| |::|   :
Db    303 VARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQILQTI 362

Qy    375 LHNLSKLQHAPSVPFQERPPDTET--------------------PEVDEDQEDGDKRWDP 414
            ||  | |||||   : | |  |                     ||   :  |||   |
Db    363 FENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDNDK 422

Qy    415 DSDMDV 420
          :||:::
Db    423 ESDVEI 428
```

RESULT 7
S22284
transcription regulator RPD3 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein NO305; protein YNL330c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 23-Mar-2001
C;Accession: S22284; S51336; S51285; S55859; S57393; S63313; S63311
R;Vidal, M.; Gaber, R.F.
Mol. Cell. Biol. 11, 6317-6327, 1991
A;Title: RPD3 encodes a second factor required to achieve maximum positive and negative
A;Reference number: S22284; MUID:92049362
A;Accession: S22284
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-433 <VID>
A;Cross-references: GB:S66438; NID:g238961; PIDN:AAB20328.1; PID:g238962
R;van Dyck, L.; Pascual-Ahuir, A.; Goffeau, A.
submitted to the EMBL Data Library, December 1994
A;Description: A 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes
icated gene for a putative aryl-alcohol dehydrogenase.
A;Reference number: S51334
A;Accession: S51336
A;Molecule type: DNA
A;Residues: 1-433 <VAN>
A;Cross-references: EMBL:X83226; NID:g642335; PIDN:CAA58228.1; PID:g642338
R;Nicaud, J.J.
submitted to the EMBL Data Library, January 1995
A;Description: Sequence analysis of a 13.9 Kb fragment of yeast chromosome XIV identifie
A;Reference number: S51285
A;Accession: S51285
A;Molecule type: DNA
A;Residues: 1-201 <NIC>
A;Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86368.1; PID:g854536
R;Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
Yeast 11, 567-572, 1995
A;Title: Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV identifies th
A;Reference number: S55859; MUID:95373280
A;Accession: S55859
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-201 <MAF>
A;Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86368.1; PID:g854536
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R;van Dyck, L.; Pascual-Ahuir, A.; Purnelle, B.; Goffeau, A.
Yeast 11, 987-991, 1995
A;Title: An 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes as we
d gene for a putative aryl-alcohol dehydrogenase.
A;Reference number: S57391; MUID:96021610
A;Accession: S57393
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-433 <VAW>
A;Cross-references: EMBL:X83226; NID:g642335; PIDN:CAA58228.1; PID:g642338
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994
R;Coster, F.; Jonniaux, J.L.; Goffeau, A.; Purnelle, B.; Van Dyck, L.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63151
A;Accession: S63313
A;Molecule type: DNA
A;Residues: 1-433 <COS>
A;Cross-references: EMBL:Z71606; NID:g1302450; PIDN:CAA96263.1; PID:g1302451; GSPDB:G
A;Experimental source: strain S288C
R;Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63287
A;Accession: S63311
A;Molecule type: DNA
A;Residues: 1-201 <MAW>
A;Cross-references: EMBL:Z71606; GSPDB:GN00014; MIPS:YNL330c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:RPD3; SDI2; SDS6; MIPS:YNL330c
A;Cross-references: SGD:S0005274; MIPS:YNL330c
A;Map position: 14L
C;Function:
A;Description: transcription regulation
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: nucleus; transcription regulation
F;23-322/Domain: RPD3/acuC homology <RAH1>

  Query Match   49.3%; Score 1341; DB 1; Length 433;
  Best Local Similarity   55.6%; Pred. No. 4.3e-96;
  Matches 230; Conservative   86; Mismatches   92; Indels   6; Gaps   2;

```
Qy     16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
          |:| |||| :|||| || |||||||||| |:|: :||| : |:: :  || ::::|:|
Db     18 KRRVAYFYDADVGNYAYGAGHPMKPHRIRMAHSLIMNYGLYKKMEIYRAKPATKQEMCQF 77

Qy     76 HADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHG 135
          | |:|: ||  :||:  :   |:  :||||:|||||||| :|    |||: |: :|| |
Db     78 HTDEYIDFLSRVTPDNLEMFKRESVKFNVGDDCPVFDGLYEYCSISGGGSMEGAARLNRG 137

Qy    136 LCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAF 195
          |:|:|:|||||||||| |||||||:||||| |:|||: | ||||:|||:||||||||||
Db    138 KCDVAVNYAGGLHHAKKSEASGFCYLNDIVLGIIELLRYHPRVLYIDIDVHHGDGVEEAF 197

Qy    196 YATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
          | |||||| ||||:|::||||| ::||| |:|| |::|||| ||||| :|  :|:|:: |
Db    198 YTTDRVMTCSFHKYGEFFPGTGELRDIGVGAGKNYAVNVPLRDGIDDATYRSVFEPVIKK 257

Qy    256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
          :|| ::| ||||||| ||||||||||||||::||| || :::|| :|::::|||||:|| 
Db    258 IMEWYQPSAVVLQCGGDSLSGDRLGCFNLSMEGHANCVNYVKSFGIPMMVVGGGGYTMRN 317

Qy    316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLL 375
          ||| ||:|||:   | ::  :| :|||||:|||| | | |||| | |: : |:::  ::
Db    318 VARTWCFETGLLNNVVLDKDLPYNEYYEYYGPDYKLSVRPSNMFNVNTPEYLDKVMTNIF 377

Qy    376 HNLSKLQHAPSVPFQERPPDTE-TPEVDEDQED-----GDKRWDPDSDMDVDDD 423
          ||   ::|||     | | |   :|:||  :     |  ::  |  :: |::
Db    378 ANLENTKYAPSVQLNHTPRDAEDLGDVEEDSAEAKDTKGGSQYARDLHVEHDNE 431
```

RESULT 8
T40300
histone deacetylase (EC 3.5.1.-) - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: cryptic loci regulator protein clr6
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T40300; T43796
R;Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duest
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21919
A;Accession: T40300
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 1-405 <LYN>
A;Cross-references: EMBL:AL023589; PIDN:CAA19053.1; GSPDB:GN00067; SPDB:SPBC36.05c
A;Experimental source: strain 972h-; cosmid c36
R;Grewal, S.I.; Bonaduce, M.J.; Klar, A.J.
Genetics 150, 563-576, 1998
A;Title: Histone deacetylase homologs regulate epigenetic inheritance of transcriptional
A;Reference number: Z22685; MUID:98429513
A;Accession: T43796
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-405 <GRE>
A;Cross-references: EMBL:AF064206; NID:g4159996; PIDN:AAD05211.1; PID:g4159997
A;Experimental source: strain 972
C;Genetics:
A;Gene: clr6; SPBC36.05c
A;Map position: 2
A;Introns: 14/1; 34/3; 51/3; 91/1; 101/2; 116/1; 131/2; 209/3; 240/3; 277/1; 324/1; 360/
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase; transcription regulation

Query Match 49.1%; Score 1336; DB 2; Length 405;
Best Local Similarity 61.7%; Pred. No. 9.5e-96;
Matches 235; Conservative 62; Mismatches 82; Indels 2; Gaps 2;

```
Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           |:|| |||| :|||:|| |||||||:|| | |: :| | : : |: | |   |: |
Db       5 KKKVSYFYDEDVGNYHYGPQHPMKPHRVRMVHNLVVNYNLYEKLNVITPVRATRNDMTRC 64

Qy      76 HADDYVSFLRSITPETQQD-QIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNH 134
           | |:|: ||  :||:| :  |  ||| ||||:|||||||| ||   ||||:| : :||
Db      65 HTDEYIEFLWRVTPDTMEKFQPHQLK-FNVGDDCPVFDGLYEFCSISAGGSIGAAQELNS 123

Qy     135 GLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEA 194
           |  :|||||||||||||| |||||||||| || ||||| |:||||:|||:|||||||
Db     124 GNAEIAINWAGGLHHAKKREASGFCYVNDIALAALELLKYHQRVLYIDIDVHHGDGVEEF 183

Qy     195 FYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMG 254
           || |||||| ||||||:|||||||:| | |:|| |::|||| |||||||  :|||::
Db     184 FYTTDRVMTCSFHKFGEYFPGTGHIKDTGIGTGKNYAVNVPLRDGIDDESYESVFKPVIS 243

Qy     255 KVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIR 314
            :|: ||| ||:|||| |||:|||||||||||:|||: || |::|||:|:: :|||||:|
Db     244 HIMQWFRPEAVILQCGTDSLAGDRLGCFNLSMKGHSMCVDFVKSFNLPMICVGGGGYTVR 303

Qy     315 NVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDL 374
           |||| | ||||:  | |::: :| ::| :|:|||| |:|  :|||| |:|| |: | :::
Db     304 NVARVWTYETGLLAGEELDENLPYNDYLQYYGPDYKLNVLSNNMENHNTRQYLDSITSEI 363

Qy     375 LHNLSKLQHAPSVPFQERPPD 395
           : ||  |  ||||   : | |
Db     364 IENLRNLSFAPSVQMHKTPGD 384
```

RESULT 9
T23963
hypothetical protein R06C1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T23963; T24071
R;Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19824
A;Accession: T23963
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-465 <WIL>
A;Cross-references: EMBL:Z81106; PIDN:CAB03224.1; GSPDB:GN00019; CESP:R06C1.1
A;Experimental source: clone R06C1
R;Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19837
A;Accession: T24071
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-465 <WI2>
A;Cross-references: EMBL:Z81108; PIDN:CAB03240.1; GSPDB:GN00019; CESP:R06C1.1
A;Experimental source: clone R09B3
C;Genetics:
A;Gene: CESP:R06C1.1
A;Map position: 1
A;Introns: 53/3; 212/1; 279/1; 326/1
C;Superfamily: RPD3 protein; RPD3/acuC homology

Query Match 49.0%; Score 1334; DB 2; Length 465;
Best Local Similarity 55.9%; Pred. No. 1.6e-95;
Matches 246; Conservative 74; Mismatches 106; Indels 14; Gaps 5;

```
Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           | :| |:|| : ||:|||||||||||||:||||:|: :||| : : |::|  |   :: |:
Db       7 KSRVSYYYDGDFGNFYYGQGHPMKPHRVRMTHSLIVNYGLYRKLNVMRPARASFSEITRY 66

Qy      76 HADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHG 135
           |:|||::|||::  :       |: ||:||||||||||:| |||   |||:  : :||
Db      67 HSDDYINFLRNVKSDNMSTFTDQMARFSVGEDCPVFDGMYEFCQLSCGGSLAAAARLNRQ 126

Qy     136 LCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAF 195
           :||||| |||||||| ||||||| |||||||||||| |:||||:|||:|||||||||
Db     127 ESEIAINWMGGLHHAKKSEASGFCYSNDIVLAILELLKHHKRVLYIDIDVHHGDGVEEAF 186

Qy     196 YATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
           | |||||||||||| |:|||||| ::|:|| |||||:|||||  ||:|| :|  :|: |||:
Db     187 YTTDRVMTVSFHKHGEYFPGTGDLKDVGAGSGKYYALNVPLRDGVDDVTYERIFRTIMGE 246

Qy     256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
           ||  |:| ||||||||||:|||||  |||:  || :||::|:|||||||:|||||||||
Db     247 VMARFQPEAVVLQCGADSLAGDRLGVFNLTTYGHGKCVEYMKSFNVPLLLVGGGGYTIRN 306

Qy     316 VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAP-SNMENKNSRQMLEEIRNDL 374
           |:||| ||| :||  || | :| |:|::|| ||| ||: | : : | |: : :::    |
Db     307 VSRCWLYETAIALNQEVSDDLPLHDYFDYFIPDYKLHIKPLAALSNFNTPEFIDQTIVAL 366

Qy     375 LHNLSKLQHAPSVPFQ--ERPPDTETPEVDE----DQEDGDKRWDP-DSDMDVDDDRKPI 427
           | || :| | ||| |       |:     ||    | :: | |    : |: |:|
Db     367 LENLKQLPHVPSVQMQSISTSCDSIVKTFDEKLIRDHQNDDVRVTQFEEDVQVED----- 421

Qy     428 PSRVKREAVEPDTKDKDGLK 447
           |    :  ||:||:   :|
Db     422 -SAEFYDGQEPETKNIQSMK 440
```

RESULT 10
T47443
hypothetical protein T18B22.80 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
C;Accession: T47443
R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24467
A;Accession: T47443
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-419 <JOR>
A;Cross-references: EMBL:AL138652
A;Experimental source: cultivar Columbia; BAC clone T18B22
C;Genetics:
A;Map position: 3
A;Introns: 47/3; 87/1; 112/1; 185/2; 206/3; 237/3; 270/1; 286/3; 317/1; 340/3; 373/3;
A;Note: T18B22.80
C;Superfamily: RPD3 protein; RPD3/acuC homology

Query Match 47.7%; Score 1298; DB 2; Length 419;

Best Local Similarity 57.0%; Pred. No. 8.7e-93;
Matches 233; Conservative 74; Mismatches 94; Indels 8; Gaps 3;

```
Qy      19 VCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFHAD 78
           |  ::   :||: |:|  ||||||||: ||| |:  |||   |:| :|  |   :: :||:
Db       4 VFFWVAGDVGSVYFGPNHPMKPHRLCMTHHLILAYGLHSKMEVYRPHKAYPIEMAQFHSP 63

Qy      79 DYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLCD 138
           ||| ||: | || |     :: |:|:|||||||: |: ||| ||||::  : :||: |||
Db      64 DYVEFLQRINPENQNLFPNEMARYNLGEDCPVFEDLFEFCQLYAGGTIDAARRLNNKLCD 123

Qy     139 IAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFYAT 198
           ||||||||||||||:|||||:||:|| |||||| | ||||:|||:|||||||||| |
Db     124 IAINWAGGLHHAKKCDASGFCYINDLVLGILELLKHHPRVLYIDIDVHHGDGVEEAFYFT 183

Qy     199 DRVMTVSFHKFGD-YFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVM 257
           |||||||||||| :|||| :::||   ||:|::|||| |||| |:: ||:     |:
Db     184 DRVMTVSFHKFGDKFFPGTGDVKEIGEREGKFYAINVPLKDGIDDSSFNRLFR----TVV 239

Qy     258 EIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRNVA 317
           ||::|||:|||||||||: |||||||||| |||||||:: ||:|||: |||||  |||
Db     240 EIYQPGAIVLQCGADSLARDRLGCFNLSIDGHAECVKFVKKFNLPLLVTGGGGYTKENVA 299

Qy     318 RCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLLHN 377
           |||  |||: |  |: :::|||::| :|| ||::| :   ::||| |::  :  |:  :| |
Db     300 RCWTVETGILLDTELPNEIPENDYIKYFAPDFSLKIPGGHIENLNTKSYISSIKVQILEN 359

Qy     378 LSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDP---DSDMDVDDD 423
           |  :|||||  || |||   |: |||::: | | |   |  :  ||:
Db     360 LRYIQHAPSVQMQEVPPDFYIPDFDEDEQNPDVRADQRSRDKQIQRDDE 408
```

RESULT 11
T11643
histone deacetylase (EC 3.5.1.-) - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: Phd1 protein
C;Species: Schizosaccharomyces pombe
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T11643; T43300
R;Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1997
A;Reference number: Z17304
A;Accession: T11643
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-434 <WOO>
A;Cross-references: EMBL:AL021046
R;Kim, Y.B.; Honda, A.; Yoshida, M.; Horinouchi, S.
FEBS Lett. 436, 193-196, 1998
A;Title: Phd1+, a histone deacetylase gene of Schizosaccharomyces pombe, is required for
A;Reference number: Z22404; MUID:98452937
A;Accession: T43300
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-434 <KIM>
A;Cross-references: EMBL:AB008888; NID:g2641698; PIDN:BAA23598.1; PID:g2641699
C;Genetics:
A;Gene: hda1; Phd1; SPAC3G9.07c
A;Map position: 1L
C;Function:
A;Description: involved in the meiotic cell cycle
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase
F;30-330/Domain: RPD3/acuC homology <RAC>

Query Match 42.9%; Score 1167.5; DB 2; Length 434;
Best Local Similarity 53.9%; Pred. No. 1.2e-82;
Matches 220; Conservative 65; Mismatches 120; Indels 3; Gaps 3;

```
Qy      16 KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRF 75
           |::| |  | :||||:|| |||||||| :|: |:  ||| |  | |  |  |   ::  |
Db      25 KKRVTYHLDEQVGNYHYGDKHPMKPHRITITNHLVMGYGLHNKMSVFSPRMATFGEMSEF 84

Qy      76 HADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHG 135
           | :||: ||: :||:  :    : ::||:|:|||||| | | |  || |:  | ||  |
Db      85 HREDYLDFLKRVTPDNAEQFADKFQQFNIGDDCPVFDGTYEFSQRSAGASLDASRKLVQG 144

Qy     136 LCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAF 195
           ||||||:|||||||: |||||||||||||| :|:   ||||:|||||||||::||
Db     145 QTDIAINWSGGLHHAKRGEASGFCYVNDIVLAILNMLRFFPRVLYIDIDIHHGDGVQQAF 204

Qy     196 YATDRVMTVSFHKF-GDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMG 254
           | :|||:|||||: ||:|| ||:  : |  |||::|||||:||| || |  ||| |:
Db     205 YESDRVLTVSFHKYNGDFFPATGNFDENGVKGGKYFALNVPLEDGIGDEQYTSLFKSIIE 264

Qy     255 KVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIR 314
           :  |:| |:|||||||||  |||| |||||  | |||:| ||||:|:|:::|||||:|
Db     265 PTINTFQPSAIVLQCGADSLGYDRLGVFNLSIHAHGECVRFTRSFNIPMLVVGGGGYTLR 324

Qy     315 NVARCWCYETGVALGVEVEDKMP-EHEYYEYFGPDYTLHV-APSNMENKNSRQMLEEIRN 372
           |||| ||||| : :  ::  ::| |  |||:| ||||||    : :||||: : ||::|
Db     325 NVARAWCYETSICVNEQIPSELPRETLYYEFFAPDYTLHPRLTTKIENKNTPKALEDLRI 384

Qy     373 DLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           |  |  |  ||||  |: |||      :|||:   |:  |   |: |
Db     385 RALEQLRYLGGAPSVQMQQIPPDLTGHLEEEDERLNDEYLDKAVDVRV 432
```

RESULT 12
T19067
hypothetical protein C08B11.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T19067
R;Sims, M.
submitted to the EMBL Data Library, November 1994
A;Reference number: Z19069
A;Accession: T19067
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-507 <WIL>
A;Cross-references: EMBL:Z46676; PIDN:CAA86662.1; GSPDB:GN00020; CESP:C08B11.2
A;Experimental source: clone C08B11
C;Genetics:
A;Gene: CESP:C08B11.2
A;Map position: 2
A;Introns: 45/1; 298/2
C;Superfamily: RPD3 protein; RPD3/acuC homology

Query Match 39.6%; Score 1077; DB 2; Length 507;
Best Local Similarity 46.7%; Pred. No. 1.5e-75;
Matches 228; Conservative 77; Mismatches 165; Indels 18; Gaps 9;

```
Qy      12 PDGV---KRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPAR 68
           |||    || | |:|  :||:::|||| ||||| |: : : |:  | : ::| |::
Db      21 PDGADVKKRNVAYYYHKDVGHFHYGQLHPMKPQRLVVCNDLVVSYEMPKYMTVVESPKLD 80

Qy      69 ERDLCRFHADDYVSFLRSITPE---TQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGS 125
           |:  || :|||:||:::||:   |  | :  |::||:|||||:| ||: :|  |||||
Db      81 AADISVFHTEDYVNFLQTVTPKLGLTMPDDV--LRQFNIGEDCPIFAGLWDYCTLYAGGS 138

Qy     126 VGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDI 185
           | |: :||| : || ||| |||||||| |||||||||||| |||||| |:||||:||||
Db     139 VEGARRLNHKMNDIVINWPGGLHHAKKSEASGFCYVNDIVLGILELLKYHKRVLYIDIDI 198

Qy     186 HHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESY 245
           ||||||:|||  :||||||||||:|| |||||:| | | | | |||:::|||||   | || |
Db     199 HHGDGVQEAFNNSDRVMTVSFHRFGQYFPGSGSIMDKGVGPGKYFAINVPLMAAIRDEPY 258

Qy     246 HLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLL 305
           ||: ::  | | | | |:|||||:|||  |||| | ||   ||  ||:::|   ||::
Db     259 LKLFESVISGVEENFNPEAIVLQCGSDSLCEDRLGQFALSFNAHARAVKYVKSLGKPLMV 318

Qy     306 LGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYY-EYFGPDYTLHVAPSNMENKNSR 364
```

```
Db    319 LGGGGYTLRNVARCWALETGVILGLRMDDEIPGTSLYSHYFTPRLLRPNLVPKMNDANSA 378

Qy    365 QMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDE-DQEDGDKRWDPDSDMDVDDD 423
Db    379 AYLASIEKETLACLRMIRGAPSVQMQ----NIVGIRLDEIEQIEENERLQKSSKSSIEYE 434

Qy    424 RKPIPSRVKREA-VEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEEASVK 482
Db    435 VGKVSEKMEEECFVEEDSKPPSFPPGQDPRRIGQYWGYDRSGLAPPRSHSDV-IEEA--K 491

Qy    483 MEEEGTNK 490
Db    492 YEDRDRRK 499


RESULT  13
S64211
probable transcription regulator YGL194c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G1330
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
C;Accession: S64211; S62051
R;Bruschi, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64183
A;Accession: S64211
A;Molecule type: DNA
A;Residues: 1-452 <BRU>
A;Cross-references: EMBL:Z72716; NID:g1322818; PIDN:CAA96906.1; PID:g1322819; MIPS:YGL19
A;Experimental source: strain S288C
R;Coglievina, M.; Delneri, D.; Zaccaria, P.; Klima, R.; Bertani, I.; Bruschi, C.V.
submitted to the EMBL Data Library, September 1995
A;Description: A 6.7 Kb fragment from chromosome VII of Saccharomyces cerevisiae contain
A;Reference number: S62051
A;Accession: S62051
A;Molecule type: DNA
A;Residues: 1-351,'HSGTHS',358,'RIIHFI',365-452 <COG>
A;Cross-references: EMBL:X91837; NID:g1177627; PIDN:CAA62950.1; PID:g1177634
A;Experimental source: strain FY1679
C;Genetics:
A;Gene: SGD:HOS2
A;Cross-references: SGD:S0003162; MIPS:YGL194c
A;Map position: 7L
C;Superfamily: RPD3 protein; RPD3/acuC homology
F;30-331/Domain: RPD3/acuC homology <RAH1>


  Query Match             39.2%;  Score 1066.5;  DB 2;  Length 452;
  Best Local Similarity   48.9%;  Pred. No. 8.1e-75;
  Matches  204;  Conservative   84;  Mismatches  120;  Indels    9;  Gaps    5;

Qy     18 KVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFHA 77
Db     27 RVSYHFNSKVSHYHYGVKHPMKPFRLMLTDHLVSSYGLHKIMDLYETRSATRDELLQFHS 86

Qy     78 DDYVSFLRSITPETQQDQIR-QLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGL 136
Db     87 EDYVNFLSKVSPENANKLPRGTLENFNIGDDCPIFQNLYDYTTLYTGASLDATRKLINNQ 146

Qy    137 CDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFY 196
Db    147 SDIAINWSGGLHHAKKNSPSGFCYVNDIVLSILNLLRYHPRILYIDIDLHHGDGVQEAFY 206

Qy    197 ATDRVMTVSFHKF-GDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGK 255
Db    207 TTDRVFTLSFHKYNGEFFPGTGDLTEIGCDKGKHFALNVPLEDGIDDDSYINLFKSIVDP 266

Qy    256 VMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRN 315
Db    267 LIMTFKPTLIVQQCGADSLGHDRLGCFNLNIKAHGECVKFVKSFGLPMLVVGGGGYTPRN 326

Qy    316 VARCWCYETGVALGVEVEDKMPEH-EYYEYFGPDYTLHVAPSNM-ENKNSRQMLEEIRND 373
Db    327 VSRLWTYETGILNDVLLPEDIPEDIPFRDSFGPDYSLYPMLDDLYENKNSKKLLEDIRIR 386

Qy    374 LLHNLSKLQHAPSVPFQERPPDTE-----TPEVDEDQEDGDKRWDPDSDMDVDDDRK 425
Db    387 CLENIRYLQGAPSVRMDAECIPTQDISALTEEEDKIIQEMNEETEADSSNRLEEMEK 443


RESULT  14
D70388
acetoin utilization protein - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 31-Mar-2000
C;Accession: D70388
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666
A;Accession: D70388
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-375 <AQF>
A;Cross-references: GB:AE000719; NID:g2983517; PIDN:AAC07100.1; PID:g2983524; GB:AE00
A;Experimental source: strain VF5
C;Genetics:
A;Gene: acuC1
C;Superfamily: acetylpolyamine aminohydrolase; RPD3/acuC homology
F;6-306/Domain: RPD3/acuC homology <RAH1>


  Query Match             17.5%;  Score 475;  DB 2;  Length 375;
  Best Local Similarity   32.4%;  Pred. No. 3.9e-29;
  Matches  116;  Conservative   67;  Mismatches  127;  Indels   48;  Gaps   10;

Qy     26 EVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFHADDYVSFLR 85
Db     11 DYGKYRYPKNHPLKIPRVSLLLRFLDAMNLIDEKELIKSRPATKEELLLFHTEDYINTLM 70

Qy     86 SITPETQQDQI---RQLKRFNV-GEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLCDIAI 141
Db     71 ----EAERCQCVPKGAREKYNIGGYENPVSYAMFTGSSLATGSTVQAIEEFLKG--NVAF 124

Qy    142 NWAGGLHHAKKCEASGFCYVNDIVLAILELLKQ-HERVLYVDIDIHHGDGVEEAFYATDR 200
Db    125 NPAGGMHHAFKSRANGFCYINDPAVGIEYLRKKGFKRILYIDLDAHHCDGVQEAFYDTDQ 184

Qy    201 VMTVSFHKFGDY-FP-GTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVME 258
Db    185 VFVLSLHQSPEYAFPFEKGFLEEIGEGKGKGYNLNIPLPKGLNDNEFLFALEKSLEIVKE 244

Qy    259 IFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPL------LLLGGGGYT 312
Db    245 VFEPEVYLLQLGTDPLLEDYLSKFNLS------NVAFLKAFNIVREVFGEGVYLGGGGYH 298

Qy    313 IRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEI 370
Db    299 PYALARAWTLIWCELSGREVPEKL----------------------NNKAKELLKSI 333


RESULT  15
S54089
hypothetical protein YPR068c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YP9499.23c
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 06-Feb-1998
C;Accession: S54089
R;Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54059
A;Accession: S54089
A;Molecule type: DNA
A;Residues: 1-470 <BAD>
```

```
A;Cross-references: EMBL:Z49219; NID:g805025; PID:g805048; MIPS:YPR068c
A;Experimental source: strain AB972
C;Genetics:
A;Gene: SGD:HOS1
A;Cross-references: SGD:S0006272; MIPS:YPR068c
A;Map position: 16R
C;Superfamily: RPD3/acuC homology
F;9-383/Domain: RPD3/acuC homology <RAH1>


  Query Match             17.3%;  Score 471.5;  DB 2;  Length 470;
  Best Local Similarity   28.6%;  Pred. No. 9.9e-29;
  Matches  128;  Conservative   65;  Mismatches  133;  Indels  121;  Gaps   14;

Qy      37 PMKPH-RIRMTHALLAHYGLLQHMQVLKPFP-ARERDLCRFHADDYVSFL---------- 84
           |   | : ::|::|:  | ||||   :  || ||: ||  ||:  |: :|
Db      20 PCNNHQKSQLTYSLINAYDLLQHFDEVLTFPYARKDDLLEFHSKSYIDYLINGRFNKMMA 79

Qy      85 ----------------------------------------------RSITPETQQDQ 95
                                                          |   |   |
Db      80 QDVNNPMVESKWSELSELADNWNEKIDYNPSQDLQRFTTRENLYNYYLNHSQALENNMDC 139

Qy      96 IR-----------------QLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLC 137
           |                  : |::|:  |||:|  |  :||   | ::     |:
Db     140 INNSEVPTNDKPTDTYILNSETKQYNLEGDCPIFSYLPMYCQVITGATLNLLDHLSPTER 199

Qy     138 DIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLK-QHERVLYVDIDIHHGDGVEEAFY 196
           | ||| || ||| |  ||||||:||:|| |  | | :  :: ||| |:||||||:||
Db     200 LIGINWDGGRHHAFKQRASGFCYINDVVLLIQRLRKAKLNKITYVDFDLHHGDGVEKAFQ 259

Qy     197 ATDRVMTVSFHKF-GDYFPGTGHIQDIGYGSGKYYS-LNVPLDDGIDDESYHLLFKPIMG 254
           : :: |:| | :   :|||||: |    | |  : :|:||  | ||    |:   |:
Db     260 YSKQIQTISVHLYEPGFFPGTGSLSD----SRKDKNVVNIPLKHGCDDNYLELIASKIVN 315

Qy     255 KVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAE-CVKFMRSF-NVPLLLLGGGGYT 312
           ::|   | |::::|| | | |||   : |:|:| :   :  |:|:    : |||||||
Db     316 PLIERHEPEALIIECGGDGLLGDRFNEWQLTIRGLSRIIINIMKSYPRAHIFLLGGGGYN 375

Qy     313 IRNVARCWCYET----------------------GVALGVEVEDKMPEHEYYEYFGPD- 348
              ::| : | |                      |  | : |  : ||:  | :  :
Db     376 DLLMSRFYTYLTWCVTKQFSNLRCGDNNSFQIDPFDVCDGDDSEQFIREHDLVEMYNEEN 435

Qy     349 --YTLHVAPSNMENKNSRQMLEEIRND 373
             | ::     ||  :  :||   |||
Db     436 YQYWIY----EMEGSSRMKML---RND 455


Search completed: April 26, 2002, 17:52:45
Job time: 108 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:50:52 ; Search time 21.38 Seconds
(without alignments)
527.322 Million cell updates/sec

Title: US-09-645-337-2
Perfect score: 2722
Sequence: 1 MDTGGNSLASGPDGVKRKVC.....KMEEEGTNKGGAEQAFPPKT 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2033.5	74.7	517	4	US-09-282-305-8	Sequence 8, Appli
2	1950.5	71.7	439	4	US-09-282-305-6	Sequence 6, Appli
3	1548	56.9	458	4	US-09-282-305-2	Sequence 2, Appli
4	1530	56.2	482	1	US-08-528-255A-1	Sequence 1, Appli
5	1530	56.2	482	1	US-08-717-365-1	Sequence 1, Appli
6	1360	50.0	432	4	US-09-282-305-10	Sequence 10, Appl
7	1130.5	41.5	351	4	US-09-282-305-4	Sequence 4, Appli
8	253.5	9.3	335	4	US-09-446-504-80	Sequence 80, Appl
9	106	3.9	754	4	US-09-214-564A-2	Sequence 2, Appli
10	105	3.9	240	2	US-08-760-745-3	Sequence 3, Appli
11	96	3.5	1435	2	US-08-568-459A-4	Sequence 4, Appli
12	96	3.5	1435	2	US-08-487-826B-4	Sequence 4, Appli
13	94.5	3.5	723	1	US-07-814-964-11	Sequence 11, Appl
14	94.5	3.5	723	1	US-08-258-442-11	Sequence 11, Appl
15	94.5	3.5	723	1	US-08-328-809-6	Sequence 6, Appli
16	94.5	3.5	723	5	PCT-US92-11107-11	Sequence 11, Appl
17	90.5	3.3	532	1	US-08-285-440-5	Sequence 5, Appli
18	90.5	3.3	532	1	US-08-630-349-5	Sequence 5, Appli
19	89.5	3.3	235	2	US-08-760-745-1	Sequence 1, Appli
20	89.5	3.3	352	1	US-08-552-142A-4	Sequence 4, Appli
21	89.5	3.3	352	1	US-08-910-973-4	Sequence 4, Appli
22	89.5	3.3	352	5	PCT-US95-05741-4	Sequence 4, Appli
23	88.5	3.3	674	1	US-08-803-973-7	Sequence 7, Appli
24	88.5	3.3	674	1	US-08-803-972-7	Sequence 7, Appli
25	88.5	3.3	707	1	US-08-803-973-12	Sequence 12, Appl
26	88.5	3.3	707	1	US-08-803-972-12	Sequence 12, Appl
27	88.5	3.3	1456	1	US-08-803-973-2	Sequence 2, Appli
28	88.5	3.3	1456	1	US-08-803-972-2	Sequence 2, Appli
29	88.5	3.3	3052	2	US-08-557-122A-26	Sequence 26, Appl
30	88	3.2	262	1	US-08-106-981-6	Sequence 6, Appli
31	88	3.2	411	2	US-08-741-134-6	Sequence 6, Appli
32	87.5	3.2	258	4	US-08-961-083-90	Sequence 90, Appl
33	87.5	3.2	565	4	US-08-961-083-218	Sequence 218, App
34	87	3.2	412	2	US-08-741-134-2	Sequence 2, Appli
35	87	3.2	1388	2	US-08-685-576-4	Sequence 4, Appli
36	86.5	3.2	237	2	US-08-760-745-5	Sequence 5, Appli
37	86.5	3.2	255	4	US-09-040-485-2	Sequence 2, Appli
38	86.5	3.2	764	1	US-08-375-300-4	Sequence 4, Appli
39	86.5	3.2	764	3	US-09-177-431-4	Sequence 4, Appli
40	86.5	3.2	764	5	PCT-US95-16930-4	Sequence 4, Appli
41	86.5	3.2	1089	1	US-08-375-300-2	Sequence 2, Appli
42	86.5	3.2	1089	3	US-09-177-431-2	Sequence 2, Appli
43	86.5	3.2	1089	5	PCT-US95-16930-2	Sequence 2, Appli
44	86	3.2	554	2	US-08-524-051-2	Sequence 2, Appli
45	86	3.2	554	3	US-09-052-778-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT   1
US-09-282-305-8
; Sequence 8, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 517
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-8


  Query Match            74.7%;  Score 2033.5;  DB 4;  Length 517;
  Best Local Similarity  76.8%;  Pred. No. 2.2e-207;
  Matches  385;  Conservative   41;  Mismatches   60;  Indels   15;  Gaps    6;

Qy       3 TGGNSLAS-GPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
           :|||||  | |||| ||:|||||||:|||||||||||||||||||||:||| ||||  |||
Db       8 SGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPMKPHRIRMTHSLLARYGLLNQMQV 67

Qy      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTY 121
             :| |||:||||||||||::||||:|||||||||| ||||||||||||||||||||||
Db      68 YRPNPARDRDLCRFHADDYINFLRSVTPETQQDQIRLLKRFNVGEDCPVFDGLYSFCQTY 127

Qy     122 AGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYV 181
           || ||||:|||||  ||||||:|||||||||||||||||||||||||| |||||||
Db     128 AGASVGGAVKLNHG-HDIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKHHERVLYV 186

Qy     182 DIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGID 241
           |||||||||||||||| ||||||||||||||||||| |:|||:  ||||||||||||||
Db     187 DIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGHSKGKYYSLNVPLDDGID 246

Qy     242 DESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNV 301
           ||||  ||||||||||:|||||||||||||||||||||||||||||||||::||||||
Db     247 DESYQSLFKPIMGKVMEVFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYMRSFNV 306

Qy     302 PLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENK 361
           |||||||||||||||||||||||||||| | ||||| :|||||||||||||||||||||
Db     307 PLLLLGGGGYTIRNVARCWCYETGVALGQEPEDKMPVNEYYEYFGPDYTLHVAPSNMENK 366
```

```
Qy    362 NSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVD 421
          |:|| |::||: || |||||:|||||| |||| |||| || ||||:| |:| ||||||:||
Db    367 NTRQQLDDIRSKLLDNLSKLRHAPSVHFQERVPDTEIPEQDEDQDDPDERHDPDSDMEVD 426

Qy    422 DD--------RKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK---VTG 470
          |         |  :  ::|||  |  |: :|| : :    :| |   :: ||:|
Db    427 DHKAVEESSRRSILGIKIKREFGENATRVQDGGR-VASEHRGLEPMAEDIGSSKQAPQAD 485

Qy    471 VNPVGVEEAS-VKMEEEGTNK 490
           : : ::| | || | | : |
Db    486 ASAMAIDEPSNVKNEPESSTK 506


RESULT   2
US-09-282-305-6
; Sequence 6, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 439
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-6


  Query Match             71.7%;  Score 1950.5;  DB 4;  Length 439;
  Best Local Similarity   84.2%;  Pred. No. 1.1e-198;
  Matches  358;  Conservative   29;  Mismatches   35;  Indels    3;  Gaps    3;

Qy      4 GGNSL-ASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVL 62
          |||||  :| || ||:|||||| |||||||||||||||||||||||||  ||||  |||
Db      8 GGNSLPTTGADGSKRRVCYFYDAEVGNYYYGQGHPMKPHRIRMTHALLGRYGLLDQMQVF 67

Qy     63 KPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYA 122
          :| |||:||||||||||||||||:|||||||||| ||||||||||||||||||||||||
Db     68 RPHPARDRDLCRFHADDYVSFLRSVTPETQQDQIRALKRFNVGEDCPVFDGLYSFCQTYA 127

Qy    123 GGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVD 182
          |||||||:||||||  |||||||||||||||||||||||||||||||||||||||| |:||||||
Db    128 GGSVGGAVKLNHG-HDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKYHQRVLYVD 186

Qy    183 IDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDD 242
          |||||||||||||| |||||||||||||||||| ||:|:|:  ||||||||||||||||
Db    187 IDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDVGHSKGKYYSLNVPLDDGIDD 246

Qy    243 ESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVP 302
          |||  |||||||||||:| ||||||||||||||||||||||||||||||||:||||||||
Db    247 ESYQSLFKPIMGKVMEVFNPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRFMRSFNVP 306

Qy    303 LLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKN 362
          |||||||||||||||||||||||||| |: |||| :|||||||||||||||||||||||
Db    307 LLLLGGGGYTIRNVARCWCYETGVALGHELTDKMPPNEYYEYFGPDYTLHVAPSNMENKN 366

Qy    363 SRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDD 422
          :|  |::|:: || |||||:||||| |||||: | || |||:|: |:| | |||::: :
Db    367 TRHQLDDIKSKLLDNLSKLRHAPSVQFQERPPEAELPEQDEDKENPDERHDADSDVEM-N 425

Qy    423 DRKPI 427
          | ||:
Db    426 DAKPL 430


RESULT   3
US-09-282-305-2
; Sequence 2, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 458
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-2


  Query Match             56.9%;  Score 1548;  DB 4;  Length 458;
  Best Local Similarity   63.3%;  Pred. No. 7.6e-156;
  Matches  280;  Conservative   66;  Mismatches   80;  Indels   16;  Gaps    3;

Qy      7 SLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFP 66
          | | | |  :|:| |||:| :|:|||||||||||||||| |:|: |||| : ::: :|:|
Db     12 SPAGGEDAHRRRVSYFYEPSIGDYYYGQGHPMKPHRIRMAHSLVVHYGLHRLLELSRPYP 71

Qy     67 ARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSV 126
          | | |: |||:||||:|| | |         | :|||||||||||||: |||  ||||:
Db     72 ASEADIRRFHSDDYVAFLASATGNPGVLDPRAIKRFNVGEDCPVFDGLFPFCQASAGGSI 131

Qy    127 GGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIH 186
          | :|||| |  || :|||||||||||| |||||||||||||||||||| | ||||||||:|
Db    132 GAAVKLNRGDADITVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKFHRRVLYVDIDVH 191

Qy    187 HGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYH 246
          |||||||||: |:|||||||||:||:|||||||  |:|   ||:|:|||||  ||||| ::
Db    192 HGDGVEEAFFTTNRVMTVSFHKYGDFFPGTGHITDVGAAEGKHYALNVPLSDGIDDTTFR 251

Qy    247 LLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLL 306
           ||: |: ||||:::|  ||||||||||:|||||||||||:||||:|::|:||:|||:::|
Db    252 GLFQCIIKKVMEVYQPDVVVLQCGADSLAGDRLGCFNLSVKGHADCLRFLRSYNVPMMVL 311

Qy    307 GGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQM 366
          |||||||||||||||| ||:||| ::|:| ::||||||||||||: | ::|| |: :
Db    312 GGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNDYYEYFGPDYTLHIQPKSVENLNTTKD 371

Qy    367 LEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKP 426
          || |:| :| ||||::| ||  | :|| | | ||  | :|| ||             | |
Db    372 LENIKNMILENLSKIEHVPSTQFHDRPSDPEAPE--EKEEDMDK------------RPP 416

Qy    427 IPSRV-KREAVEPDTKDKDGLK 447
            ||:    | : ||:| | ||
Db    417 QRSRLWSGGAYDSDTEDPDSLK 438


RESULT   4
US-08-528-255A-1
; Sequence 1, Application US/08528255A
; Patent No. 5659016
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  : FURUKAWA, YOICHI
;    TITLE OF INVENTION:  RPDL PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
```

```
;       STREET:  2026 Rambling Road
;       CITY:  Kalamazoo
;       STATE:  Michigan
;       COUNTRY:  USA
;       ZIP:  49008-1699
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;       COMPUTER:  IBM PC/XT/AT Compatible
;       OPERATING SYSTEM:  MS-DOS 5.0
;       SOFTWARE:  WordPerfect 5.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/528,255A
;       FILING DATE:  September 14, 1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP6-227876
;       FILING DATE:  22-SEPTEMBER-1994
;       APPLICATION NUMBER:  JP7-183763
;       FILING DATE:  20-JULY-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Terryence F. Chapman
;       REGISTRATION NUMBER:  32549
;       REFERENCE/DOCKET NUMBER:  Furuya Case 1335
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (616) 381-1156
;       TELEFAX:  (616) 381-5465
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  482
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     IMMEDIATE SOURCE:
;       LIBRARY:  human fetal lung cDNA library
US-08-528-255A-1


  Query Match             56.2%;  Score 1530;  DB 1;  Length 482;
  Best Local Similarity   59.2%;  Pred. No. 6.8e-154;
  Matches  290;  Conservative   67;  Mismatches   99;  Indels   34;  Gaps    7;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | :|||||:|| :|||||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 GTRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::|||||||||||||: |||   |||| :||||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVKLN 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
               |||:|||||||||| ||||||||||||||||||| |:||||:|||||||||||
Db     126 KQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGVEE 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| |||||||||||:|:|||||| ::||| | ||||::| || |||||||  :|||:|
Db     186 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVM 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           ||||:|:| |||||||:|||||||||||||:|||||:||:|::|||:|:|:||||||||
Db     246 SKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGYTI 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           |||||| ||| ||| |: :::| ::|:||||||: ||::|||| |:|: : ||:|:
Db     306 RNVARCRTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIKQR 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD--------------PDSDM 418
           |  ||  | ||| |  |  | |    |  |||::| |||                ||:
Db     366 LFENLRMLPHAPGVQMQAIPEDAIPEESGDEDEDDPDKRISICSSDKRIACEEEFSDSEE 425

Qy     419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
           :  :  ||   :     ||| |    | |:||      |  |    || |   ||
Db     426 EGEGGRKNSSNFKKAKRVKTE----DEKEKD-----PEEKK----EVTEEEKTKEEKPEA 472

Qy     474 VGVEEASVKM 483
           ||:|  ||:
Db     473 KGVKE-EVKL 481


RESULT   5
US-08-717-365-1
; Sequence 1, Application US/08717365
; Patent No. 5763182
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  :  FURUKAWA, YOICHI
;    TITLE OF INVENTION:  RPDL PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,365
;      FILING DATE:  23-SEP-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/528 255
;      FILING DATE:  September 14, 1995
;      APPLICATION NUMBER:  JP6-227876
;      FILING DATE:  22-SEPTEMBER-1994
;      APPLICATION NUMBER:  JP7-183763
;      FILING DATE:  20-JULY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1335
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  482
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  human fetal lung cDNA library
US-08-717-365-1


  Query Match             56.2%;  Score 1530;  DB 1;  Length 482;
  Best Local Similarity   59.2%;  Pred. No. 6.8e-154;
  Matches  290;  Conservative   67;  Mismatches   99;  Indels   34;  Gaps    7;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | :|||||:|| :|||||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 GTRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::|||||||||||||: |||   |||| :||||
Db      66 KYHSDDYIKFLRSIRPDNMSEYSKQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVKLN 125
```

```
Qy    134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
          |||:|||||||||||| ||||||||||||||||||| |:||||:|||||||||||||
Db    126 KQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGVEE 185

Qy    194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
          ||| |||||||||||:|:|||||| ::||| | |||||:| || |||||||| :|||:|
Db    186 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKPVM 245

Qy    254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
          |||||:|:| |||||||:||||||||||||:|||||:||:|:::|||:|:|:|||||||
Db    246 SKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGYTI 305

Qy    314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
          |||||| ||| ||| |: :::| ::|:|||||: ||::|||| |:|: : ||:|:
Db    306 RNVARCRTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIKQR 365

Qy    374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD----------PDSDM 418
          |  ||  | ||| |  |  | |    |  |||::| |||              ||:
Db    366 LFENLRMLPHAPGVQMQAIPEDAIPEESGDEDEDDPDKRISICSSDKRIACEEEEFSDSEE 425

Qy    419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
          : :  ||   :     ||| |    | |:||      |  |    || |   ||
Db    426 EGEGGRKNSSNFKKAKRVKTE----DEKEKD-----PEEKK----EVTEEEKTKEEKPEA 472

Qy    474 VGVEEASVKM 483
          ||:|  ||:
Db    473 KGVKE-EVKL 481


RESULT   6
US-09-282-305-10
; Sequence 10, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 432
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-10


  Query Match             50.0%;  Score 1360;  DB 4;  Length 432;
  Best Local Similarity   59.0%;  Pred. No. 6.8e-136;
  Matches  249;  Conservative   61;  Mismatches   92;  Indels   20;  Gaps    3;

Qy     21 YFYDP---EVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFHA 77
          ||  |   :||| |:|  |||||||: ||| |:  ||| | |:: :|  |   :| :|||:
Db      8 YFITPIVGDVGNVYFGPNHPMKPHRLCMTHHLVLSYGLHQKMEIYRPHKAYPIELAQFHS 67

Qy     78 DDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLC 137
          ||| ||  |||::|     :| |:|:|||||||| |: ||| ||||::  : :||| :|
Db     68 ADYVEFLHRITPDSQHLYASELTRYNLGEDCPVFDNLFEFCQIYAGGTLDAARRLNHKIC 127

Qy    138 DIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFYA 197
          ||||||||||||||||||||||:||:|| |||||| | ||||:|||:||||||||||
Db    128 DIAINWAGGLHHAKKCEASGFCYINDLVLGILELLKYHARVLYIDIDVHHGDGVEEAFYF 187

Qy    198 TDRVMTVSFHKFGD-YFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKV 256
          |||||||||||:|| :||||| |:|||   ||||::|:|| ||||| |:   || |: ||
Db    188 TDRVMTVSFHKYGDLFFPGTGDIKDIGEREGKYYAINIPLKDGIDDTSFTRPFKTIIAKV 247

Qy    257 MEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRNV 316
          :| : |||:|||||||||: ||||||||||:||||||:: ||:|||: ||||||  ||
Db    248 VETYLPGAIVLQCGADSLARDRLGCFNLSIEGHAECVKFVKKFNIPLLVTGGGGYTKENV 307

Qy    317 ARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLLH 376
          ||||  |||| |  |: :::|::|| ||| ||||| |   ||:| ||:  |  |:  ::
Db    308 ARCWAVETGVLLDTELPNEIPKNEYIEYFAPDYTLKVPNLNMDNLNSKTYLSSIKVQVME 367

Qy    377 NLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKR----------WDPDSDMDV 420
          :|  :|||| |  || |||   |: |||: | |:|              :: |:| |
Db    368 SLRYIQHAPGVQMQEVPPDFYIPDFDEDELDPDERVDQHTQDKQIHRDDEYYEGDNDNDH 427

Qy    421 DD 422
          ||
Db    428 DD 429


RESULT   7
US-09-282-305-4
; Sequence 4, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 351
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-4


  Query Match             41.5%;  Score 1130.5;  DB 4;  Length 351;
  Best Local Similarity   48.9%;  Pred. No. 1.2e-111;
  Matches  216;  Conservative   50;  Mismatches   53;  Indels  123;  Gaps    4;

Qy      7 SLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFP 66
          | | | |  :|:| |||:| :|:||||||||||
Db     12 SPAGGEDAHRRRVSYFYEPSIGDYYYGQGHPMKP---------------------------- 45

Qy     67 ARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSV 126

Db     46 ------------------------------------------------------------ 45

Qy    127 GGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIH 186
                              ||||| |||||||||||||||||| | ||||||||:|
Db     46 ----------------------HHAKKSEASGFCYVNDIVLAILELLKFHRRVLYVDIDVH 84

Qy    187 HGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYH 246
          |||||||||: |:|||||||||:|||:|||||| |:|   ||:|:||||| ||||| ::
Db     85 HGDGVEEAFFTTNRVMTVSFHKYGDFFPGTGHITDVGAAEGKHYALNVPLSDGIDDTTFR 144

Qy    247 LLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLL 306
            ||: |: ||||:::|  |||||||||:|||||||||:||||:|:::|:||:|||:::|
Db    145 GLFQCIIKKVMEVYQPDVVVLQCGADSLAGDRLGCFNLSVKGHADCLRFLRSYNVPMMVL 204

Qy    307 GGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQM 366
          |||||||||||||||| ||:||| ::|:| ::|||||||||||: | ::|| |: :
Db    205 GGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNDYYEYFGPDYTLHIQPKSVENLNTTKD 264

Qy    367 LEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKP 426
          || |:| :| ||||::| ||  | :|| | | ||  | :|| ||             | |
Db    265 LENIKNMILENLSKIEHVPSTQFHDRPSDPEAPE--EKEEDMDK------------RPP 309

Qy    427 IPSRV-KREAVEPDTKDKDGLK 447
```

```
                ||:    | : ||:| | ||
Db    310 QRSRLWSGGAYDSDTEDPDSLK 331


RESULT   8
US-09-446-504-80
; Sequence 80, Application US/09446504
; Patent No. 6218150
; GENERAL INFORMATION:
;  APPLICANT:  UEMORI, Takashi
;  APPLICANT:  SATO, Yoshimi
;  APPLICANT:  FUJITA, Tomoko
;  APPLICANT:  MIYAKE, Kazue
;  APPLICANT:  MUKAI, Hiroyuki
;  APPLICANT:  ASADA, Kiyozo
;  APPLICANT:  KATO, Ikunoshin
;  TITLE OF INVENTION:  DNA POLYMERASE-RELATED FACTORS
;  FILE REFERENCE:  1422-408PCT
;  CURRENT APPLICATION NUMBER: US/09/446,504
;  CURRENT FILING DATE:  1999-12-23
;  PRIOR APPLICATION NUMBER:  PCT/JP98/02845
;  PRIOR FILING DATE:  1998-06-24
;  PRIOR APPLICATION NUMBER:  JP 9-187496
;  PRIOR FILING DATE:  1997-06-26
;  PRIOR APPLICATION NUMBER:  JP 9-320692
;  PRIOR FILING DATE:  1997-11-27
;  NUMBER OF SEQ ID NOS:  92
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 80
;   LENGTH: 335
;   TYPE: PRT
;   ORGANISM: Pyrococcus furiosus
US-09-446-504-80


  Query Match             9.3%;  Score 253.5;  DB 4;  Length 335;
  Best Local Similarity   24.2%;  Pred. No. 2.1e-18;
  Matches   88;  Conservative   71;  Mismatches  155;  Indels   49;  Gaps   13;

Qy      22 FYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFHADDYV 81
           ||  :   :   : ||  | |: :  :  :   || : ::  :| |  |  : : |  ||:
Db       6 FYSKKFNLHRPSEYHPENPKRLEIVLSKVRELGLEERIE--EPNPVEETFVEKIHDRDYI 63

Qy      82 SFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVK-LNHGLCDIA 140
           :|::    |  :  | :|       |  |  | :|      | :   ::  |::|   :|
Db      64 NFVK----EAVEKGITRL-----DPDTYVSPGTWSAALLALGAARSAALSALHYGGLHMA 114

Qy     141 INWAGGLHHAKKCEAS-----GFCYVNDIVLAILELLKQH-ERVLYVDIDIHHGDGVEEA 194
           :    | |  ::  |     |||  |:   |:: | ::   :|: :| | |||:| :|
Db     115 LVRPPGHHAGRRGRAMGAPTLGFCIFNNAASAVVTLKEEGVGKVVVIDFDAHHGNGTQEI 174

Qy     195 FYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYS-LNVPLDDGIDDESYHLLFKPIM 253
           |:    |: :  |:  | :||:| : ::| ||  | | :|:|:     |  |  ::  |:
Db     175 FWNDPDVIHIDLHE-RDIYPGSGDVSEVG-GSNAYGSKINLPMPHYSGDGDYIYVWDEIV 232

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
             ::|  :|  :|:  | |   || |    |:  ::     :  : :  :| || |  :
Db     233 LPIVEEVKPKVIVISAGFDGFKGDGLTTLRLTESFYSYAGATLNKYPLAFILEGGYGVGL 292

Qy     314 RN--VARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIR 371
           |    || | |       |  |   ||      || :|             ||::
Db     293 DKGFPAFIMGYEEGKA-------KAREEPRYE------TLKLA------------EEVK 326

Qy     372 NDL 374
           : |
Db     327 DIL 329


RESULT   9
US-09-214-564A-2
; Sequence 2, Application US/09214564A
; Patent No. 6150515
; GENERAL INFORMATION:
;  APPLICANT:  Sharp, Phillip A.
;  APPLICANT:  Zhou, Qiang
;  TITLE OF INVENTION:  TAT-SF: Cofactor For Stimulation Of Transcriptional
;  TITLE OF INVENTION:  Elongation By HIV-1 TAT
;  FILE REFERENCE:  M0656/7042
;  CURRENT APPLICATION NUMBER: US/09/214,564A
;  CURRENT FILING DATE:  1999-08-18
;  PRIOR APPLICATION NUMBER:  US 60/021,218
;  PRIOR FILING DATE:  1996-07-03
;  PRIOR APPLICATION NUMBER:  US 60/033,152
;  PRIOR FILING DATE:  1996-12-13
;  PRIOR APPLICATION NUMBER:  PCT/US97/11713
;  PRIOR FILING DATE:  1997-07-03
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 754
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-214-564A-2


  Query Match             3.9%;  Score 106;  DB 4;  Length 754;
  Best Local Similarity   20.4%;  Pred. No. 0.04;
  Matches   80;  Conservative   35;  Mismatches  129;  Indels  148;  Gaps   15;

Qy     230 YSLNVPLDDGIDDESYHLLFKPIMGKVMEIFRPGAV------VLQCGADSLSGDRLGCF- 282
           |   :| |  :|:       |  :| |   | |          : :    :| || | |:
Db     136 YVSGLPPDITVDE------FIQLMSKFGIIMRDPQTEEFKVKLYKDNQGNLKGDGLCCYL 189

Qy     283 -----NLSI----------KGHAECVKFMRSFNVPLLLLGGGGYTIRNVARCWCYETGV 326
                 |::           | | |  ||          | |     :   :|  |:  :
Db     190 KRESVELALKLLDEDEIRGYKLHVEVAKFQ---------LKGEYDASKKKKKCKDYKKKL 240

Qy     327 ALGVEVEDKMPEHEYYEYFGPDYTLH---VAPSN----MENKNSRQMLEEIRNDLLHNLS 379
           ::  :  |  ||        ||    |   |    |   |: ::   :| ||| ||    |
Db     241 SMQQKQLDWRPERR----AGPSRMRHERVVIIKNMFHPMDFEDDPLVLNEIREDLRVECS 296

Qy     380 KLQHAPSVPFQERPPDTETPEVDEDQEDGD-----------------KRWDPDSDMDVD 421
           |      :   :| ||        | |: |                 : ||  :|  |:
Db     297 KFGQIRKLLLFDRHPDGVASVSFRDPEEADYCIQTLDGRWFGGRQITAQAWDGTTDYQVE 356

Qy     422 DD--------------------------------RKPIPSRVKREAVEPDT--------- 440
           :                                ||  ||| :  :  | |
Db     357 ETSREREERLRGWEAFLNAPEANRGLSVQILSLLRKAGPSRARHFSEHPSTSKMNAQETA 416

Qy     441 ---------------KDKDGLKGIMERG----------------KGCEVEVDESGSTKVT 469
                         | :||  |  | |                :||  :  | |  |
Db     417 TGMAFEEPIDEKKFEKTEDG--GEFEEGASENNAKESSPEKEAEEGCPEKESEEGCPK-- 472

Qy     470 GVNPVGVE-EASVKMEEEGTNKGGAEQAFPPK 500
                 | |   | |  |||    |:|:  | |
Db     473 ----RGFEGSCSQKESEEGNPVRGSEEDSPKK 500


RESULT  10
US-08-760-745-3
; Sequence 3, Application US/08760745
; Patent No. 5972658
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Murry, Lynn E.
;    TITLE OF INVENTION:  LUNG GROWTH FACTOR VARIANT
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94304
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/760,745
;      FILING DATE:  Herewith
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0169 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  240 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  598956
US-08-760-745-3


  Query Match             3.9%;  Score 105;  DB 2;  Length 240;
  Best Local Similarity   23.8%;  Pred. No. 0.0075;
  Matches   48;  Conservative   27;  Mismatches   77;  Indels   50;  Gaps    7;

Qy     334 DKMPE-------HEYYE-YFGPDYTLHVAPSNM-------------ENKNSRQMLEEIR 371
           |:|||       ::|   :||   |  : | ::              : |   : | ||
Db      31 DEMPEAAVKSTANKYQVFFFGTHETAFLGPKDLFPYEESKEKFGKPNKRKGFSEGLWEIE 90

Qy     372 NDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMD------VDDDRK 425
           |:     |  | :      | |     || :  : ||||: : :   |      :|:  |
Db      91 NNPTVKASGYQSSQKKSCVEEP----EPEPEAAEGDGDKKGNAEGSSDEEGKLVIDEPAK 146

Qy     426 ----------------PIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKV 468
                           | | | ||  |: ::|:     :||    ||| : : |
Db     147 EKNEKGALKRRAGDLLEDSPKRPK-EAENPEGEEKEAATLEVERPLPMEVEKNSTPSEPG 205

Qy     469 TGVNPVGVEEASVKMEEEGTNK 490
           :|  |   ||     ||| | :
Db     206 SGRGPPQEEEEEDEEEEATKE 227


RESULT  11
US-08-568-459A-4
; Sequence 4, Application US/08568459A
; Patent No. 5849306
;  GENERAL INFORMATION:
;    APPLICANT:  Sim, Kim L.
;    APPLICANT:  Chitnis, Chetan
;    APPLICANT:  Miller, Louis H.
;    APPLICANT:  Peterson, David S.
;    APPLICANT:  Su, Xin-zhaun
;    APPLICANT:  Wellems, Thomas E.
;    TITLE OF INVENTION:  BINDING DOMAINS FROM PLASMODIUM VIVAX
;    TITLE OF INVENTION:  AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe Martens Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/568,459A
;      FILING DATE:  07-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH121.001CP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1435 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Plasmodium falciparum
US-08-568-459A-4


  Query Match             3.5%;  Score 96;  DB 2;  Length 1435;
  Best Local Similarity   18.6%;  Pred. No. 1.4;
  Matches   76;  Conservative   61;  Mismatches  155;  Indels  116;  Gaps   19;

Qy     169 LELLKQHERVLYVDIDIHHGDGVEEAFYATD-----RVMTVSFHKFGDYFPGTGHIQDIG 223
           | ::|:|     : | |:    ::  :   |     :::  :|    |  || :  |:
Db     509 LLMIKEH----ILAIAIYESRILKRKYKNKDDKEVCKIINKTFADIRDIIGGTDYWNDLS 564

Qy     224 --------YGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLS 275
                   : ||   |   |    || : :: | :   :  :|:   |   | | :
Db     565 NRKLVGKINTNSKYVHRNKKNDKLFRDEWWKVIKKDVWNVISWVFKDKTV---CKEDDIE 621

Qy     276 GDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGY---------TIR------------ 314
                         :| | |:       |  |          |::
Db     622 ---------------NIPQFFRWFSE-----WGDDYCQDKTKMIETLKVECKEKPCEDDN 661

Qy     315 NVARCWCYETGVALGVEVEDKMPEHEYYEY-FGPDYTLH-----VAPSNMENKNSR---- 364
             ::|  |:  ::   |  :|  : :| ||  | :| ::     : |     | |
Db     662 CKSKCNSYKEWISKKKEEYNKQAK-QYQEYQKGNNYKMYSEFKSIKPEVYLKKYSEKCSN 720

Qy     365 -----QMLEEIRNDLLHNLSKLQHAPSVPF----------QERPPD--------TETPEV 401
                :  ||: :|  :  :       ||          ||  |:        |||| :
Db     721 LNFEDEFKEELHSDYKNKCTMCPEVKDVPISIIRNNEQTSQEAVPEENTEIAHRTETPSI 780

Qy     402 DEDQEDGDKRWDPDSDMDVDDDRKPIPSRVKREAVEPDTKDKDGL-----------KGI 449
           |  : |::: : |     ||    |    :    | | ||   |            | :
Db     781 SEGPK-GNEQKERD-----DDSLSKISVSPENSRPETDAKDTSNLLKLKGDVDISMPKAV 834

Qy     450 MERGKGCEVEVDESGSTKVTGVN--PVGVEEASVKMEEEGTNKGGAEQ 495
           :       : | | |   ::|||  |:  :    | | |  |   :|:
Db     835 IGSSPNDNINVTEQGD-NISGVNSKPLSDDVRPDKKELEDQNSDESEE 881


RESULT  12
US-08-487-826B-4
; Sequence 4, Application US/08487826B
; Patent No. 5993827
;  GENERAL INFORMATION:
;    APPLICANT:  Sim, Kim L.
;    APPLICANT:  Chitnis, Chetan
;    APPLICANT:  Miller, Louis H.
;    APPLICANT:  Peterson, David S.
;    APPLICANT:  Su, Xin-zhaun
;    APPLICANT:  Wellems, Thomas E.
```

```
;    TITLE OF INVENTION:  BINDING DOMAINS FROM PLASMODIUM VIVAX
;    TITLE OF INVENTION:  AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Knobbe Martens Olson & Bear
;      STREET:  620 Newport Center Drive  16th Floor
;      CITY:  Newport Beach
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  92660
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,826B
;      FILING DATE:  10-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Israelsen, Ned
;      REGISTRATION NUMBER:  29,655
;      REFERENCE/DOCKET NUMBER:  NIH121.001CP1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 235-8550
;      TELEFAX:  (619) 235-0176
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1435 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Plasmodium falciparum
US-08-487-826B-4


  Query Match             3.5%;  Score 96;  DB 2;  Length 1435;
  Best Local Similarity   18.6%;  Pred. No. 1.4;
  Matches   76;  Conservative   61;  Mismatches  155;  Indels  116;  Gaps   19;

Qy     169 LELLKQHERVLYVDIDIHHGDGVEEAFYATD-----RVMTVSFHKFGDYFPGTGHIQDIG 223
           | ::|:|     : | |:    ::  :   |     :::  :|    |   || :  |:
Db     509 LLMIKEH----ILAIAIYESRILKRKYKNKDDKEVCKIINKTFADIRDIIGGTDYWNDLS 564

Qy     224 --------YGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLS 275
                  : ||   |   |    || : :: | :   :  :|:   |   |  | :
Db     565 NRKLVGKINTNSKYVHRNKKNDKLFRDEWWKVIKKDVWNVISWVFKDKTV---CKEDDIE 621

Qy     276 GDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGY---------TIR------------ 314
                       :| | |:       |  |         |::
Db     622 ---------------NIPQFFRWFSE-----WGDDYCQDKTKMIETLKVECKEKPCEDDN 661

Qy     315 NVARCWCYETGVALGVEVEDKMPEHEYYEY-FGPDYTLH-----VAPSNMENKNSR---- 364
             ::|  |:  ::   |  :|  : :| ||  | :| ::     : |     | |
Db     662 CKSKCNSYKEWISKKKEEYNKQAK-QYQEYQKGNNYKMYSEFKSIKPEVYLKKYSEKCSN 720

Qy     365 -----QMLEEIRNDLLHNLSKLQHAPSVPF----------QERPPD--------TETPEV 401
                :  ||: :|  :  :      ||           ||  |:         |||| :
Db     721 LNFEDEFKEELHSDYKNKCTMCPEVKDVPISIIRNNEQTSQEAVPEENTEIAHRTETPSI 780

Qy     402 DEDQEDGDKRWDPDSDMDVDDDRKPIPSRVKREAVEPDTKDKDGL-----------KGI 449
           |  : |::: : |     ||    |     :    | | ||   |            | :
Db     781 SEGPK-GNEQKERD-----DDSLSKISVSPENSRPETDAKDTSNLLKLKGDVDISMPKAV 834

Qy     450 MERGKGCEVEVDESGSTKVTGVN--PVGVEEEASVKMEEEGTNKGGAEQ 495
           :       : | | |   ::|||  |:  :    | | |  |   :|:
Db     835 IGSSPNDNINVTEQGD-NISGVNSKPLSDDVRPDKKELEDQNSDESEE 881


RESULT  13
```

```
US-07-814-964-11
; Sequence 11, Application US/07814964
; Patent No. 5359047
;  GENERAL INFORMATION:
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Lippard, Stephen J.
;    TITLE OF INVENTION:  DNA Structure Specific Recognition
;    TITLE OF INVENTION:  Protein and Uses Therefor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/814,964
;      FILING DATE:  19911226
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/539,906
;      FILING DATE:  18-JUN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-4787AAA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  11
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic I
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG-box
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
US-07-814-964-11
```

```
  Query Match             3.5%;  Score 94.5;  DB 1;  Length 723;
  Best Local Similarity   20.8%;  Pred. No. 0.63;
  Matches   80;  Conservative   38;  Mismatches   92;  Indels  175;  Gaps   21;

Qy     231 SLNVPLDDGIDDESYH---LLFKP--------------------------------IMG 254
           ||: |:  |     ||   ||| |                                 :||
Db     258 SLDPPIKQG--QTRYHYLVLLFAPDEETTIELPFSEAELRDKYEGKLEKEISGPVYEVMG 315

Qy     255 KVMEIF------RPGAVVLQCGADSLSGDRLGCFNLSIKGH------------------ 289
           |||::        ||  :   |  :     :||   :  |:
Db     316 KVMKVLIGRKITGPGNFIGHSGTAA-----VGCSFKAAAGYLYPLERGFIYIHKPPLHIR 370

Qy     290 ---AECVKFMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALG-VEVEDKMPEHEYYEYF 345
               | | ||          || |     | : :|  :  | | :   : : || : |
Db     371 FEEISSVNFARS----------GGST-----RSFDFEVTLKNGTVHIFSSIEKEEYAKLF 415

Qy     346 GPDY----TLHVAPSNM-ENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPD----- 395
             ||     |||  ||| ::|:  : ::   :|                 |  ||
Db     416 --DYITQKKLHV--SNMGKDKSGYKDVDFGDSD---------------NENEPDAYLAR 455

Qy     396 --TETPEVDEDQEDGD-------------------KRWDPDSDMDVDDD---------- 423
             |  | :|| :|||                     : :| : : | |||
Db     456 LKAEAREKEEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGGDS 515

Qy     424 ----RKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK------VTGVNP 473
               :|      |:|  |   |:|       || |    :  :|| |      :  :|
Db     516 DGAKKKKEKKSEKKEKKEKKHKEK-------ERTKKPSKKKKDSGKPKRATTAFMLWLND 568

Qy     474 VGVEEASVKMEEEG------TNKGG 492
               |:| |  |         |||
Db     569 T---RESIKRENPGIKVTEIAKKGG 590


RESULT  14
US-08-258-442-11
; Sequence 11, Application US/08258442
; Patent No. 5670621
;  GENERAL INFORMATION:
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Lippard, Stephen J.
;    TITLE OF INVENTION:  DNA Structure Specific Recognition
;    TITLE OF INVENTION:  Protein and Uses Therefor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  2 Militia Drive
;      CITY:  Lexington
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/258,442
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/539,906
;      FILING DATE:  18-JUN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Granahan, Patricia
;      REGISTRATION NUMBER:  32,227
;      REFERENCE/DOCKET NUMBER:  MIT-4787AAA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-861-6240
;      TELEFAX:  617-861-9540
;  INFORMATION FOR SEQ ID NO:  11:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic I
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG-box
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
US-08-258-442-11


  Query Match             3.5%;  Score 94.5;  DB 1;  Length 723;
  Best Local Similarity   20.8%;  Pred. No. 0.63;
  Matches   80;  Conservative   38;  Mismatches   92;  Indels  175;  Gaps   21;

Qy     231 SLNVPLDDGIDDESYH---LLFKP--------------------------------IMG 254
           ||: |:  |     ||   ||| |                                 :||
Db     258 SLDPPIKQG--QTRYHYLVLLFAPDEETTIELPFSEAELRDKYEGKLEKEISGPVYEVMG 315

Qy     255 KVMEIF------RPGAVVLQCGADSLSGDRLGCFNLSIKGH------------------ 289
           |||::        ||  :   |  :     :||   :  |:
Db     316 KVMKVLIGRKITGPGNFIGHSGTAA-----VGCSFKAAAGYLYPLERGFIYIHKPPLHIR 370

Qy     290 ---AECVKFMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALG-VEVEDKMPEHEYYEYF 345
               | | ||          || |     | : :|  :  | | :   : : || : |
Db     371 FEEISSVNFARS----------GGST-----RSFDFEVTLKNGTVHIFSSIEKEEYAKLF 415

Qy     346 GPDY----TLHVAPSNM-ENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPD----- 395
             ||     |||  ||| ::|:  : ::   :|                 |  ||
Db     416 --DYITQKKLHV--SNMGKDKSGYKDVDFGDSD---------------NENEPDAYLAR 455

Qy     396 --TETPEVDEDQEDGD-------------------KRWDPDSDMDVDDD---------- 423
             |  | :|| :|||                     : :| : : | |||
Db     456 LKAEAREKEEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGGDS 515

Qy     424 ----RKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK------VTGVNP 473
               :|      |:|  |   |:|       || |    :  :|| |      :  :|
Db     516 DGAKKKKEKKSEKKEKKEKKHKEK-------ERTKKPSKKKKDSGKPKRATTAFMLWLND 568

Qy     474 VGVEEASVKMEEEG------TNKGG 492
               |:| |  |         |||
Db     569 T---RESIKRENPGIKVTEIAKKGG 590


RESULT  15
US-08-328-809-6
```

```
; Sequence 6, Application US/08328809
; Patent No. 5705334
;  GENERAL INFORMATION:
;    APPLICANT:  Lippard, Stephen J.
;    APPLICANT:  Essigmann, John M.
;    APPLICANT:  Donahue, Brian A.
;    APPLICANT:  Toney, Jeffrey H.
;    APPLICANT:  Bruhn, Suzanne L.
;    APPLICANT:  Pil, Pieter M.
;    APPLICANT:  Brown, Steven
;    APPLICANT:  Kellett, Patti
;    TITLE OF INVENTION:  Uses For DNA Structure-Specific
;    TITLE OF INVENTION:  Recognition Proteins
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patent Administrator, Testa, Hurwitz & Thibeault
;      STREET:  53 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/328,809
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fenton, Gillian M.
;      REGISTRATION NUMBER:  36,508
;      REFERENCE/DOCKET NUMBER:  MIT-023 (5473/24)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-248-7000
;      TELEFAX:  617-248-7100
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  723 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila melanogaster
;    IMMEDIATE SOURCE:
;      CLONE:  Drosophila SSRP (predicted)
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  458..507
;      OTHER INFORMATION:  /label= Acidic
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  518..547
;      OTHER INFORMATION:  /label= Basic I
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  547..620
;      OTHER INFORMATION:  /label= HMG
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  632..649
;      OTHER INFORMATION:  /label= Basic II
;    FEATURE:
;      NAME/KEY:  Domain
;      LOCATION:  657..723
;      OTHER INFORMATION:  /label= Mixed Charge
US-08-328-809-6


  Query Match             3.5%;  Score 94.5;  DB 1;  Length 723;
  Best Local Similarity   20.8%;  Pred. No. 0.63;
  Matches   80;  Conservative   38;  Mismatches   92;  Indels  175;  Gaps   21;
```

```
Qy    231 SLNVPLDDGIDDESYH---LLFKP--------------------------------IMG 254
          ||: |:  |     ||   ||| |                                 :||
Db    258 SLDPPIKQG--QTRYHYLVLLFAPDEETTIELPFSEAELRDKYEGKLEKEISGPVYEVMG 315

Qy    255 KVMEIF------RPGAVVLQCGADSLSGDRLGCFNLSIKGH------------------ 289
          |||::        ||  :   |  :     :||   :  |:
Db    316 KVMKVLIGRKITGPGNFIGHSGTAA-----VGCSFKAAAGYLYPLERGFIYIHKPPLHIR 370

Qy    290 ---AECVKFMRSFNVPLLLLGGGGYTIRNVARCWCYETGVALG-VEVEDKMPEHEYYEYF 345
              | | ||          || |     | : :|  :  | | :   : : || : |
Db    371 FEEISSVNFARS----------GGST-----RSFDFEVTLKNGTVHIFSSIEKEEYAKLF 415

Qy    346 GPDY----TLHVAPSNM-ENKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPD----- 395
            ||      |||  ||| ::|:  : ::   :|                 |  ||
Db    416 --DYITQKKLHV--SNMGKDKSGYKDVDFGDSD---------------NENEPDAYLAR 455

Qy    396 --TETPEVDEDQEDGD-------------------KRWDPDSDMDVDDD---------- 423
            |  | :|| :|||                    : :| : : | |||
Db    456 LKAEAAREKEEDDDDGDSDEESTDEDFKPNENESDVAEEYDSNVESDSDDDSDASGGGGDS 515

Qy    424 ----RKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK------VTGVNP 473
              :|      |:|  |   |:|       || |     :  :||  |      :  :|
Db    516 DGAKKKKEKKSEKKEKKEKKHKEK-------ERTKKPSKKKKDSGKPKRATTAFMLWLND 568

Qy    474 VGVEEEASVKMEEEG------TNKGG 492
            |:| |  |         |||
Db    569 T---RESIKRENPGIKVTEIAKKGG 590


Search completed: April 26, 2002, 17:51:24
Job time: 32 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:50:52 ; Search time 43 Seconds
(without alignments)
863.040 Million cell updates/sec

Title: US-09-645-337-2
Perfect score: 2722
Sequence: 1 MDTGGNSLASGPDGVKRKVC.....KMEEEGTNKGGAEQAFPPKT 501

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2722	100.0	501	22	AAB67811	Amino acid sequenc
2	2722	100.0	501	22	AAB67165	Arabidopsis histon
3	2138	78.5	473	21	AAY58829	Soybean histone de
4	2033.5	74.7	517	20	AAY28800	Maize histone deac
5	2005	73.7	493	21	AAY58828	Rice histone deace
6	1950.5	71.7	439	20	AAY28799	Maize histone deac
7	1568.5	57.6	471	21	AAG05187	Arabidopsis thalia
8	1568.5	57.6	471	21	AAG39372	Arabidopsis thalia
9	1560.5	57.3	471	22	AAB67812	Amino acid sequenc
10	1548	56.9	458	20	AAY28797	Maize histone deac
11	1544	56.7	482	18	AAW29324	A novel histone de
12	1544	56.7	482	22	AAB67164	Human histone deac
13	1544	56.7	482	22	AAB49954	Human histone deac
14	1544	56.7	488	21	AAB56985	Human prostate can
15	1530	56.2	482	17	AAR88919	RPDL, a human foet
16	1494	54.9	488	22	AAB49955	Human histone deac
17	1421.5	52.2	432	21	AAG05188	Arabidopsis thalia
18	1421.5	52.2	432	21	AAG39373	Arabidopsis thalia
19	1412	51.9	370	22	AAU00243	Human histone deac
20	1382.5	50.8	425	21	AAG05189	Arabidopsis thalia
21	1382.5	50.8	425	21	AAG39374	Arabidopsis thalia
22	1373	50.4	428	22	AAB49956	Human histone deac
23	1360	50.0	432	20	AAY28801	Maize histone deac
24	1341	49.3	433	22	AAB67166	Yeast histone deac
25	1130.5	41.5	351	20	AAY28798	Maize histone deac
26	855.5	31.4	415	22	AAM40374	Human polypeptide
27	852	31.3	377	21	AAY83027	Human histone deac
28	852	31.3	377	22	AAB84361	Amino acid sequenc
29	629	23.1	180	22	AAB70496	Saccharomyces cere
30	566.5	20.8	296	22	AAM42160	Human polypeptide
31	494.5	18.2	120	21	AAY58830	Wheat histone deac
32	483	17.7	133	18	AAW29329	A partial deacetyl
33	477	17.5	375	22	AAU00241	Cys75Ser/Cys77Ser
34	475	17.5	375	22	AAU00240	Wild-type histone
35	471	17.3	375	22	AAU00242	Tyr297Phe mutant o
36	456	16.8	125	18	AAW29330	A partial deacetyl
37	333.5	12.3	398	21	AAG07246	Arabidopsis thalia
38	328.5	12.1	468	21	AAG51592	Arabidopsis thalia
39	328.5	12.1	499	21	AAG51591	Arabidopsis thalia
40	328.5	12.1	552	21	AAG51590	Arabidopsis thalia
41	327.5	12.0	364	21	AAG07248	Arabidopsis thalia
42	327.5	12.0	377	21	AAG07247	Arabidopsis thalia
43	321	11.8	69	18	AAW29326	HDx protein family
44	320.5	11.8	468	21	AAG09532	Arabidopsis thalia
45	320.5	11.8	499	21	AAG09531	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAB67811
ID AAB67811 standard; Protein; 501 AA.
XX
AC AAB67811;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a histone deacetylase designated AtRPD3A.
XX
KW Histone deacetylase; AtRPD3A; RPD3; gene expression; transgenic plant;
KW HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS Arabidopsis thaliana.
XX
PN CA2316036-A1.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX
DR WPI; 2001-258457/27.
DR N-PSDB; AAF80350.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT -

```
XX
PS   Claim 10; Fig 1A; 91pp; English.
XX
CC   The present sequence represents Arabidopsis thaliana histone deacetylase
CC   designated AtRPD3A. The protein is homologous to yeast RPD3 and HDA1.
CC   The polynucleotide sequence is used in the method of the invention.
CC   The specification describes a method for regulating gene expression in
CC   transgenic plants. The method comprises modifying histones by introducing
CC   chimeric nucleotide sequences which have regulatory elements in operative
CC   association with a gene of interest or with a nucleotide sequence
CC   encoding histone deacetylase. The method is useful for regulating the
CC   developmental, physiological or biochemical pathway within a plant,
CC   particularly for repressing ethylene-responsive phenotypes
CC   (e.g. inhibition of hypocotyl elongation). The method is also useful
CC   as a functional test for identifying a phenotype associated with
CC   perturbing a gene. The histone deacetylase genes are useful for
CC   altering the development of an organism.
XX
SQ   Sequence   501 AA;


  Query Match             100.0%;  Score 2722;  DB 22;  Length 501;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-275;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mdtggnslasgpdgvkrkvcyfydpevgnyyygqghpmkphrirmthallahygllqhmq 60

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 vlkpfparerdlcrfhaddyvsflrsitpetqqdqirqlkrfnvgedcpvfdglysfcqt 120

Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 yaggsvggsvklnhglcdiainwagglhhakkceasgfcyvndivlailellkqhervly 180

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 vdidihhgdgveeafyatdrvmtvsfhkfgdyfpgtghiqdigygsgkyyslnvpldDgi 240

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ddesyhllfkpimgkvmeifrpgavvlqcgadslsgdrlgcfnlsikghaecvkfmrsfn 300

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vpllllggggytirnvarcwcyetgvalgvevedkmpeheyyeyfgpdytlhvapsnmen 360

Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 knsrqmleeirndllhnlsklqhapsvpfqerppdtetpevdedqedgdkrwdpdsdmdv 420

Qy     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 dddrkpipsrvkreavepdtkdkdglkgimergkgcevevdesgstkvtgvnpvgveeas 480

Qy     481 VKMEEEGTNKGGAEQAFPPKT 501
           |||||||||||||||||||||
Db     481 vkmeeegtnkggaeqafppkt 501


RESULT   2
AAB67165
ID   AAB67165 standard; Protein; 501 AA.
XX
AC   AAB67165;
XX
DT   12-APR-2001  (first entry)
XX
DE   Arabidopsis histone deacetylase 1 HDAC1.
XX
KW   Gene expression; disease model; treatment; suppression;
```

```
KW   chromatin inactivation; DNA binding site; histone deacetylation.
XX
OS   Arabidopsis thaliana.
XX
PN   WO200102019-A2.
XX
PD   11-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-GB02497.
XX
PR   30-JUN-1999;   99GB-0015126.
XX
PA   (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI   Buluwela L,  Ali S;
XX
DR   WPI; 2001-138069/14.
XX
PT   Suppressing expression of selected gene for treating cancer, involves
PT   introducing peptide comprising nucleic acid binding portion that binds
PT   to or associated with selected gene and chromatin inactivation portion
PT   -
XX
PS   Disclosure; Fig 5; 65pp; English.
XX
CC   The present invention describes a method of suppressing the expression of
CC   a gene of interest by introducing into the cell a protein containing a
CC   DNA binding site and a chromatin inactivation portion, or a nucleotide
CC   encoding such a peptide. Preferably the chromatin inactivation portion is
CC   part of the histone deacetylation (HDAC) complex. The method can be used
CC   in disease treatment, for example in the treatment of cancer by the
CC   suppression of oncogenes, and in the production of disease models.
XX
SQ   Sequence   501 AA;


  Query Match             100.0%;  Score 2722;  DB 22;  Length 501;
  Best Local Similarity   100.0%;  Pred. No. 3.7e-275;
  Matches  501;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mdtggnslasgpdgvkrkvcyfydpevgnyyygqghpmkphrirmthallahygllqhmq 60

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 vlkpfparerdlcrfhaddyvsflrsitpetqqdqirqlkrfnvgedcpvfdglysfcqt 120

Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 yaggsvggsvklnhglcdiainwagglhhakkceasgfcyvndivlailellkqhervly 180

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 vdidihhgdgveeafyatdrvmtvsfhkfgdyfpgtghiqdigygsgkyyslnvpldDgi 240

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ddesyhllfkpimgkvmeifrpgavvlqcgadslsgdrlgcfnlsikghaecvkfmrsfn 300

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 vpllllggggytirnvarcwcyetgvalgvevedkmpeheyyeyfgpdytlhvapsnmen 360

Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 knsrqmleeirndllhnlsklqhapsvpfqerppdtetpevdedqedgdkrwdpdsdmdv 420

Qy     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 dddrkpipsrvkreavepdtkdkdglkgimergkgcevevdesgstkvtgvnpvgveeas 480

Qy     481 VKMEEEGTNKGGAEQAFPPKT 501
```

```
          ||||||||||||||||||||
Db    481 vkmeeegtnkggaeqafppkt 501


RESULT   3
AAY58829
ID   AAY58829 standard; Protein; 473 AA.
XX
AC   AAY58829;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Soybean histone deacetylase 1 (HD1) protein.
XX
KW   Chromatin associated protein; histone deacetylase gene 1; HD1;
KW   soybean; transgenic plant; transcription regulation.
XX
OS   Glycine max.
XX
PN   WO200004177-A1.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15807.
XX
PR   14-JUL-1998;   98US-0092841.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Vollmer SJ;
XX
DR   WPI; 2000-182439/16.
DR   N-PSDB; AAZ58260.
XX
PT   New nucleic acid fragment useful as probes and primers, for
PT   transforming plants  -
XX
PS   Claim 1; Page 27-28; 36pp; English.
XX
CC   The present sequence is that of soybean histone deacetylase 1
CC   (HD1), a chromatin associated protein, as deduced from a soybean
CC   root cDNA clone (see AAZ58260) isolated on the basis of homology to
CC   plant histone deacetylases.  The invention relates to isolated
CC   rice, soybean and wheat nucleic acid fragments encoding HD1.  It
CC   also relates to the construction of a chimeric gene encoding all or
CC   a portion of HD1, in sense or antisense orientation, where
CC   expression of the chimeric gene results in production of altered
CC   levels of HD1 in a transformed host cell.  The availability of
CC   nucleic acid sequences encoding (portions) of histone deacetylase
CC   proteins will facilitate studies of global transcriptional
CC   regulation in eukaryotic cells, and will also provide mechanisms to
CC   control transcriptional gene regulation in plants.
XX
SQ   Sequence   473 AA;


  Query Match           78.5%;  Score 2138;  DB 21;  Length 473;
  Best Local Similarity  80.7%;  Pred. No. 3.5e-214;
  Matches  392;  Conservative   33;  Mismatches   37;  Indels   24;  Gaps    2;

Qy       1 MDTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQ 60
           |::|||| || ||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1 mesggnslpsgsdgvkrkvsyfydpevgnyyygqghpmkphrirmthallahygllqhmq 60

Qy      61 VLKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQT 120
           ||||  |::|||:|||||||:||| |||||||||:|||||||||||||||||||||||||
Db      61 vlkpmaakdrdlckfhaddyvaflrgitpetqqdqlrqlkrfnvgedcpvfdglysfcqt 120

Qy     121 YAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLY 180
           ||||||||::|||||:|||||||||||||||||||||||||||||||||||||| ||||||
Db     121 yaggsvggalklnhgvcdiainwagglhhakkceasgfcyvndivlailellkihervly 180

Qy     181 VDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGI 240
           |||||||||||||| |||||||||||||||||| |:|||| ||||||||||||
Db     181 vdidihhgdgveeafyttdrvmtvsfhkfgdyfpgtgdirdigyakgkyyslnvplddgi 240

Qy     241 DDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFN 300
           ||||| |||||||||||||||||||||||||||||||||||||||||||::|||||
Db     241 ddesyqslfkpimgkvmeifrpgavvlqcgadslsgdrlgcfnlsikghaecvrymrsfn 300

Qy     301 VPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMEN 360
           |||||||||||||||||||:|| ||||:|::||||:|||||||||||||||||||||||
Db     301 vpllllggggytirnvarcwcfetsvalgielddkmpqheyyeyfgpdytlhvapsnmen 360

Qy     361 KNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDV 420
           |||||:|:|||  || |||:|||||||||||||| |  | ||||:| |:|||||| :|
Db     361 knsrqlldeirakllднlsrlqhapsvpfqerppdaellerdedqddrderwdpdsdrev 420

Qy     421 DDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNPVGVEEAS 480
           ||  |:  ||| | |  | :|||                       |:||: : |:|
Db     421 gddsnpvrrrvksecv--daedkd----------------------tvsgvdsmavdepc 456

Qy     481 VKMEEE 486
           :| |::
Db     457 ikeeqd 462


RESULT   4
AAY28800
ID   AAY28800 standard; protein; 517 AA.
XX
AC   AAY28800;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-4.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90840.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 63-65; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   517 AA;
```

```
  Query Match             74.7%;  Score 2033.5;  DB 20;  Length 517;
  Best Local Similarity   76.8%;  Pred. No. 3.3e-203;
  Matches  385;  Conservative   41;  Mismatches   60;  Indels   15;  Gaps    6;

Qy       3 TGGNSLAS-GPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
           :|||||  | |||| ||:|||||||:|||||||||||||||||||:||| |||| |||
Db       8 sggnslpsvgpdgqkrrvcyfydpdvgnyyyggghpmkphrirmthsllarygllnqmqv 67

Qy      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTY 121
             :| |||:||||||||||::||||:|||||||||| ||||||||||||||||||||||
Db      68 yrpnpardrdlcrfhaddyinflrsvtpetqqdqirllkrfnvgedcpvfdglysfcqty 127

Qy     122 AGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYV 181
           || ||||:|||||| ||||||:|||||||||||||||||||||||||| |||||||
Db     128 agasvggavklnhg-hdiainwsgglhhakkceasgfcyvndivlailellkhhervlyv 186

Qy     182 DIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGID 241
           ||||||||||||||| |||||||||||||||||| |:|||:  ||||||||||||||
Db     187 didihhgdgveeafyttdrvmtvsfhkfgdyfpgtgdirdighskgkyyslnvplddgid 246

Qy     242 DESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNV 301
           ||||  |||||||||||:|||||||||||||||||||||||||||||||::||||||
Db     247 desyqslfkpimgkvmevfrpgavvlqcgadslsgdrlgcfnlsikghaecvrymrsfnv 306

Qy     302 PLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENK 361
           |||||||||||||||||||||||||||| | ||||| :|||||||||||||||||||||
Db     307 plllggggytirnvarcwcyetgvalgqepedkmpvneyyeyfgpdytlhvapsnmenk 366

Qy     362 NSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVD 421
           |:|| |::||: || |||||:||||| |||| |||| || ||||:| |:| ||||||:||
Db     367 ntrqqlddirsklldnlsklrhapsvhfqervpdteipeqdedqddpderhdpdsdmevd 426

Qy     422 DD--------RKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK---VTG 470
           |         |  :  ::||| |  |: :|| : :    :| |   :: ||:|
Db     427 dhkaveessrrsilgikikrefgenatrvqdggr-vasehrglepmaedigsskqapqad 485

Qy     471 VNPVGVEEAS-VKMEEEGTNK 490
             : : ::| | || | | : |
Db     486 asamaidepsnvknepesstk 506


RESULT   5
AAY58828
ID   AAY58828 standard; Protein; 493 AA.
XX
AC   AAY58828;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Rice histone deacetylase 1 (HD1) protein (C-terminal portion).
XX
KW   Chromatin associated protein; histone deacetylase gene 1; HD1;
KW   rice; transgenic plant; transcription regulation.
XX
OS   Oryza sativa.
XX
PN   WO200004177-A1.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15807.
XX
PR   14-JUL-1998;   98US-0092841.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Vollmer SJ;
XX
DR   WPI; 2000-182439/16.
DR   N-PSDB; AAZ58259.
XX
PT   New nucleic acid fragment useful as probes and primers, for
PT   transforming plants -
XX
PS   Claim 1; Page 25-26; 36pp; English.
XX
CC   The present sequence is that of a C-terminal portion of rice
CC   histone deacetylase 1 (HD1), a chromatin associated protein, as
CC   deduced from a cDNA contig (see AAZ58259) obtained from cDNA clones
CC   isolated from rice callus, 15-day-old leaf and 15-day-old seedling
CC   cDNA libraries on the basis of homology to other plant histone
CC   deacetylases.  The invention relates to isolated rice, soybean and
CC   wheat nucleic acid fragments encoding HD1.  It also relates to the
CC   construction of a chimeric gene encoding all or a portion of HD1,
CC   in sense or antisense orientation, where expression of the chimeric
CC   gene results in production of altered levels of HD1 in a
CC   transformed host cell.  The availability of nucleic acid sequences
CC   encoding (portions) of histone deacetylase proteins will facilitate
CC   studies of global transcriptional regulation in eukaryotic cells,
CC   and will also provide mechanisms to control transcriptional gene
CC   regulation in plants.
XX
SQ   Sequence   493 AA;


  Query Match             73.7%;  Score 2005;  DB 21;  Length 493;
  Best Local Similarity   78.1%;  Pred. No. 2.9e-200;
  Matches  378;  Conservative   36;  Mismatches   54;  Indels   16;  Gaps    5;

Qy      21 YFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLCRFHADDY 80
           |||| |||||||||||||||||||||||||||||  |||||| |||:|||||||||||
Db       1 yfydaevgnyyyggghpmkphrirmthallahygglldqmqvlkphpardrdlcrfhaddy 60

Qy      81 VSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLNHGLCDIA 140
           |:||||:|||||||||| ||||||||||||||||||||||||||:||||||  |||
Db      61 vaflrsvtpetqqdqiralkrfnvgedcpvfdglysfcqtyaggsvggavklnhg-hdia 119

Qy     141 INWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEEAFYATDR 200
           |||||||||||||||||||||||||||||||||| |:|||||||||||||||||| |||
Db     120 inwagglhhakkceasgfcyvndivlailellkyhqrvlyvdidihhgdgveeafyttdr 179

Qy     201 VMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIMGKVMEIF 260
           |||||||||||||||| |:|||:  ||||||||||||||||||  ||||||||||:|
Db     180 vmtvsfhkfgdyfpgtgdirdighskgkyyslnvplddgiddesyqslfkpimgkvmevf 239

Qy     261 RPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTIRNVARCW 320
           |||||||||||||||||||||||||:||||||:|||||||||||||||||||||||||||
Db     240 rpgavvlqcgadslsgdrlgcfnlsirghaecvrfmrsfnvpllllggggytirnvarcw 299

Qy     321 CYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRNDLLHNLSK 380
           ||||||||| |: |||| :||:|||||||||||||||||:|| |::||: || ||||
Db     300 cyetgvalgheltdkmppneyfeyfgpdytlhvapsnmenkntrqqlddirsrlldnlsk 359

Qy     381 LQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKPIPS----------R 430
           |:||||| ||||||: | || |||||| |:|   |||::: || ||:           |
Db     360 lrhapsvqfqerppeaelpeqdedqedpderhhadsdvem-ddvkplddsgrrssiqnvr 418

Qy     431 VKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTK---VTGVNPVGVEE-ASVKMEEE 486
           ||||: | |  |:|| :   | || |   |  ||:|    | : : ::|  |:|:| :
Db     419 vkresaetdaadqdgnrvaaentkgtepaadgvgsskqtvptdasamaidepgslkvepd 478

Qy     487 GTNK 490
             :||
Db     479 nsnk 482


RESULT   6
AAY28799
ID   AAY28799 standard; protein; 439 AA.
XX
AC   AAY28799;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-3.
```

```
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90839.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 58-60; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   439 AA;


  Query Match             71.7%;  Score 1950.5;  DB 20;  Length 439;
  Best Local Similarity   84.2%;  Pred. No. 1.2e-194;
  Matches  358;  Conservative   29;  Mismatches   35;  Indels    3;  Gaps    3;

Qy       4 GGNSL-ASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVL 62
           |||||  :| || ||:|||||| ||||||||||||||||||||||  ||||  |||
Db       8 ggnslpttgadgskrrvcyfydaevgnyyygqghpmkphrirmthallgryglldqmqvf 67

Qy      63 KPFPARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYA 122
           :| |||:|||||||||||||||:|||||||||| ||||||||||||||||||||||||
Db      68 rphpardrdlcrfhaddyvsflrsvtpetqqdqiralkrfnvgedcpvfdglysfcqtya 127

Qy     123 GGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVD 182
           ||||||:|||||| ||||||||||||||||||||||||||||||||||| |:||||||
Db     128 ggsvggavklnhg-hdiainwagglhhakkceasgfcyvndivlailellkyhqrvlyvd 186

Qy     183 IDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDD 242
           |||||||||||||| |||||||||||||||||||| |:|:|:  ||||||||||||||||
Db     187 idihhgdgveeafyttdrvmtvsfhkfgdyfpgtgdirdvghskgkyyslnvplddgidd 246

Qy     243 ESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVP 302
           |||  |||||||||||:| ||||||||||||||||||||||||||||:||||||||
Db     247 esyqslfkpimgkvmevfnpgavvlqcgadslsgdrlgcfnlsikghaecvrfmrsfnvp 306

Qy     303 LLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKN 362
           |||||||||||||||||||||||||| |: |||| :|||||||||||||||||||||||
Db     307 llllggggytirnvarcwcyetgvalgheltdkmppneyyeyfgpdytlhvapsnmenkn 366

Qy     363 SRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDD 422
           :|  |::|:: || |||||:|||| |||||||: | || |||:|: |:| | |||::: :
Db     367 trhqlddikskllldnlsklrhapsvqfqerppeaelpeqdedkenpderhdadsdvem-n 425

Qy     423 DRKPI 427
           | ||:
Db     426 dakpl 430


RESULT   7
AAG05187
ID   AAG05187 standard; Protein; 471 AA.
XX
AC   AAG05187;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 1499.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
```

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 57.6%; Score 1568.5; DB 21; Length 471;
Best Local Similarity 60.9%; Pred. No. 1.1e-154;
Matches 293; Conservative 64; Mismatches 93; Indels 31; Gaps 5;

```
Qy       2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
           |  | || ||||| ||:| |||:| :|:||||||||||||||||| |:|: || | : :::
Db       4 desgislpsgpdgrkrrvsyfyeptigdyyyyqqghpmkphrirmahsliihyhlhrrlei 63

Qy      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119
             :|  |   |: |||: :|| || |::||:  |    | |:||||||||||||: ||:
Db      64 srpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrfnvgedcpvfdglfdfcr 123
```

```
Qy    120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179
Db    124 asaggsigaavklnrqdadiainwggglhhakkseasgfcyvndivlgilellkmfkrvl 183

Qy    180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239
Db    184 yididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdvgaekgkyyalnvplndg 243

Qy    240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299
Db    244 mddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcfnlsvkghadclrflrsy 303

Qy    300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359
Db    304 nvplmvlggggytirnvarcwcyetavavgvepdnklpyneyfeyfgpdytlhvdpspme 363

Qy    360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDP----- 414
Db    364 nlntpkdmerirntlleqlsglihapsvqfqhtppvnrv--ldepeddmetrpkpriwsg 421

Qy    415 --DSDMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472
Db    422 tatyesdsdddkp------------------lhgyscrg-gattdrdstgedemdddn 461

Qy    473 P 473
Db    462 p 462
```

RESULT 8
AAG39372
ID AAG39372 standard; Protein; 471 AA.
XX
AC AAG39372;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48705.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

```
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             57.6%;  Score 1568.5;  DB 21;  Length 471;
  Best Local Similarity   60.9%;  Pred. No. 1.1e-154;
  Matches  293;  Conservative   64;  Mismatches   93;  Indels   31;  Gaps    5;

Qy       2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
           |  | || ||||| ||:| |||:| :|:|||||||||||||||| |:|: || | : :::
Db       4 desgislpsgpdgrkrrvsyfyeptigdyyygqghpmkphrirmahsliihyhlhrrlei 63

Qy      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119
            :|  |   |: ||| :|| || |::||:  |    | |:||||||||||||||: ||:
Db      64 srpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrfnvgedcpvfdglfdfcr 123

Qy     120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179
             ||||:| :|||    |||||| |||||||| ||||||||||||| ||||||  :|||
Db     124 asaggsigaavklnrqdadiainwggglhhakkseasgfcyvndivlgilellkmfkrvl 183

Qy     180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239
           |:|||:|||||||||| |||||||||||||:|||||||:|:|   ||||:|||||:||
Db     184 yididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdvgaekgkyyalnvplndg 243

Qy     240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299
           :|||||:  ||:|:|:: ||||:::| ||||||||||||||||||||:||||:|::|:||:
Db     244 mddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcfnlsvkghadclrflrsy 303

Qy     300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359
           ||||::|||||||||||||||| ||:||| ::|:| :||:|||||||||| || ||
Db     304 nvplmvlggggytirnvarcwcyetavavgvepdnklpyneyfeyfgpdytlhvdpspme 363

Qy     360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDP----- 414
           | |: : :| ||| ||  || | ||||| ||  ||      :|| ::| : |  |
Db     364 nlntpkdmerirntlleqlsglihapsvqfqhtppvnrv--ldepeddmetrpkpriwsg 421

Qy     415 --DSDMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472
             : | ||| ||                  | |   || |   : | :|  ::   |
Db     422 tatyesdsddddkp------------------lhgyscrg-gattdrdstgedemdddn 461

Qy     473 P 473
           |
Db     462 p 462


RESULT   9
AAB67812
ID   AAB67812 standard; Protein; 471 AA.
XX
AC   AAB67812;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Amino acid sequence of a histone deacetylase designated AtRPD3B.
XX
KW   Histone deacetylase; AtRPD3B; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS   Arabidopsis thaliana.
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
```

XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX
DR WPI; 2001-258457/27.
DR N-PSDB; AAF80351.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT -
XX
PS Claim 10; Fig 1B; 91pp; English.
XX
CC The present sequence represents Arabidopsis thaliana histone deacetylase
CC designated AtRPD3B. The protein is homologous to yeast RPD3 and HDA1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operative
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.
XX
SQ Sequence 471 AA;

Query Match 57.3%; Score 1560.5; DB 22; Length 471;
Best Local Similarity 60.7%; Pred. No. 7.5e-154;
Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;

```
Qy       2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61
Db       4 desgislpsgpdgrkrrvsyfyeptigdyyygqghpmkphrirmahsliihyhlhrrlei 63

Qy      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119
Db      64 srpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrfnvgedcpvfdglfdfcr 123

Qy     120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179
Db     124 asaggsigaavklnrqdadiainwggglhhakkseasgfcyvndivlgilellkmfkrvl 183

Qy     180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239
Db     184 yididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdvgaekgkyyalnvplndg 243

Qy     240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299
Db     244 mddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcfnlsvkghadclrflrsy 303

Qy     300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359
Db     304 nvplmvlggegytirnvarcwcyetavavgvepdnklpyneyfeyfgpdytlhvdpspme 363

Qy     360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDP----- 414
Db     364 nlntpkdmerirntlleqlsglihapsvqfqhtppvnrv--ldepeddmetrpkpriwsg 421

Qy     415 --DSDMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472
Db     422 tatyesdsddddkp------------------lhgyscrg-gattdrdstgedemdddn 461
```

Qy 473 P 473
Db 462 p 462

RESULT 10
AAY28797
ID AAY28797 standard; protein; 458 AA.
XX
AC AAY28797;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-1.
XX
KW Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
XX
OS Zea mays.
XX
PN WO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
DR WPI; 1999-611038/52.
DR N-PSDB; AAX90837.
XX
PT New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance -
XX
PS Claim 1; Page 49-51; 87pp; English.
XX
CC The present sequence is maize histone deacetylase encoded by HD cDNA
CC belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl
CC modifications, may be localised to promoters targeted for repression by
CC other proteins that associate with HD and specifically bind regulatory
CC elements in promoter DNA. The HD nucleotide sequence can be used for
CC producing transgenic plants with increased disease resistance.
CC Additionally, compositions find use in screening for toxins that affect
CC pathogenicity and in determining which disease response promoters are
CC regulated by histone deacetylase.
XX
SQ Sequence 458 AA;

Query Match 56.9%; Score 1548; DB 20; Length 458;
Best Local Similarity 63.3%; Pred. No. 1.4e-152;
Matches 280; Conservative 66; Mismatches 80; Indels 16; Gaps 3;

```
Qy       7 SLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFP 66
Db      12 spaggedahrrrvsyfyepsigdyyygqghpmkphrirmahslvvhyglhrllelsrpyp 71

Qy      67 ARERDLCRFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSV 126
Db      72 aseadirrfhsddyvaflasatgnpgvldpraikrfnvgedcpvfdglfpfcqasaqqsi 131

Qy     127 GGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIH 186
Db     132 gaavklnrgdaditvnwagglhhakkseasgfcyvndivlailellkfhrrvlyvdidvh 191

Qy     187 HGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYH 246
```

```
Db    192 hgdgveeaffttnrvmtvsfhkygdffpgtghitdvgaaegkhyalnvplsdgiddttfr 251

Qy    247 LLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLL 306
          ||: |: |||||:::|  ||||||||||:||||||||:||||:|::|:||:|||:::|
Db    252 glfqciikkvmevyqpdvvvlqcgadslagdrlgcfnlsvkghadclrflrsynvpmmvl 311

Qy    307 GGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQM 366
          |||||||||||||||||| ||:||| ::|:| ::|||||||||||||: | ::|| |: :
Db    312 ggggytirnvarcwcyetavavgvepdnklpyndyyeyfgpdytlhiqpksvenlnttkd 371

Qy    367 LEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDSDMDVDDDRKP 426
          || |:| :| ||||::| ||  | :|| | | || | :|| ||             | |
Db    372 leniknmilenlskiehvpstqfhdrpsdpeape--ekeedmdk------------rpp 416

Qy    427 IPSRV-KREAVEPDTKDKDGLK 447
            ||:    | : ||:| | ||
Db    417 qrsrlwsggaydsdtedpdslk 438


RESULT  11
AAW29324
ID   AAW29324 standard; Protein; 482 AA.
XX
AC   AAW29324;
XX
DT   16-APR-1998  (first entry)
XX
DE   A novel histone deacetylase (HDx) protein, designated HD1.
XX
KW   Histone deacetylase gene; HDx; HD1; HDx polypeptide; deacetylation; H3;
KW   H4; cell differentiation; chromatin structure; cell cycle progression;
KW   proliferative disorder; fibroproliferative disorder;
KW   degenerative disorder; autoimmune disease; HDx inhibitor.
XX
OS   Homo sapiens.
XX
PN   WO9735990-A2.
XX
PD   02-OCT-1997.
XX
PF   26-MAR-1997;   97WO-US05275.
XX
PR   26-MAR-1996;   96US-0624735.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Hassig CA,  Jamison TF,  Schreiber SL,  Taunton J;
XX
DR   WPI; 1997-489651/45.
DR   N-PSDB; AAT86371.
XX
PT   New isolated histone deacetylase polypeptide(s) and genes - used to
PT   develop products for modulating the proliferation, survival or
PT   differentiation of cells, e.g. for treating tumours.
XX
PS   Claim 6; Pages 112-114; 159pp; English.
XX
CC   The present sequence represents a novel histone deacetylase (HDx)
CC   polypeptide, designated HD1. The HDx polypeptides are capable of
CC   modulating proliferation survival and differentiation of cells. The
CC   proteins are able to alter chromatin structure by deacetylating histones
CC   such as H3 or H4. They have the ability to modulate cell growth by
CC   influencing cell cycle progression or to modulate gene transcription. The
CC   products can be used for diagnosis and therapy. They can be used, for
CC   example, to treat tumours or proliferative disorders or spermatogenesis,
CC   osteogenesis, chondrogenesis or the differentiation of progenitor cells.
CC   They can also be used to treat psoriasis, bone diseases,
CC   fibroproliferative disorders, degenerative disorders, or for repair of
CC   cartilage, increasing bone density, liver repair subsequent to a partial
CC   hepatectomy, to promote regeneration of lung tissue in the treatment of
CC   emphysema, or for inducing tolerance in autoimmune diseases, and
CC   transplant recipients. HDx inhibitors can be used as anti-fungal agents,
CC   preservatives in foodstuff, feed supplements for promoting weight gain in
CC   livestock, disinfectants, insecticides or defoliants.  The products can
CC   also be used in cell cultures.
CC   note: in the claims, the full length HD1 sequence is incorrectly
CC   referred to as SEQ. ID. No.2; from the disclosure it is clear that HD1
CC   is SEQ. ID. No.5.
XX
SQ   Sequence   482 AA;


  Query Match             56.7%;  Score 1544;  DB 18;  Length 482;
  Best Local Similarity   59.4%;  Pred. No. 4.1e-152;
  Matches  291;  Conservative   67;  Mismatches   98;  Indels   34;  Gaps    7;

Qy     14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
          | :|||||:|| :|||||||||||||||||||| || :||| : |:: :|  |   ::
Db      6 gtrrkvcyyydgdvgnyyygqghpmkphrirmthnlllnyglyrkmeiyrphkanaeemt 65

Qy     74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
          ::|:|||: ||||| |:   :  :|::|||||||||||||: |||   ||||  :||||
Db     66 kyhsddyikflrsirpdnmseyskqmqrfnvgedcpvfdglfefcqlstggsvasavkln 125

Qy    134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
             |||:|||||||||| |||||||||||||||||| |:||||:||||||||||||
Db    126 kqqtdiavnwagglhhakkseasgfcyvndivlailellkyhqrvlyididihhgdgvee 185

Qy    194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
          ||| ||||||||||:|:|||||| ::||| | ||||::| || ||||||||  :|||:|
Db    186 afyttdrvmtvsfhkygeyfpgtgdlrdigagkgkyyavnyplrdgiddesyeaifkpvm 245

Qy    254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           ||||:|:| |||||||:|||||||||||||:|||||:||:|::|||:|:|:|||||||
Db    246 skvmemfqpsavvlqcgsdslsgdrlgcfnltikghakcvefvksfnlpmlmlggggyti 305

Qy    314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
          ||||||| ||| |||  |: :::| ::|:|||||||: ||::|||| |:|: : ||:|:
Db    306 rnvarcwtyetavaldteipnelpyndyfeyfgpdfklhispsnmtnqntneylekikqr 365

Qy    374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD--------------PDSDM 418
          |  ||  | ||| |  |  | |    |  |||::| |||                 ||:
Db    366 lfenlrmlphapgvqmqaipedaipeesgdededdpdkrisicssdkriaceeefsdsee 425

Qy    419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
          : :  ||   :     ||| |    | |:||      |  |    || |   ||
Db    426 egeggrknssnfkkakrvkte---dekekd-----peekk---evteeektkeekpea 472

Qy    474 VGVEEASVKM 483
            ||:|  ||:
Db    473 kgvke-evkl 481


RESULT  12
AAB67164
ID   AAB67164 standard; Protein; 482 AA.
XX
AC   AAB67164;
XX
DT   12-APR-2001  (first entry)
XX
DE   Human histone deacetylase 1 HDAC1.
XX
KW   Human; gene expression; disease model; treatment; suppression;
KW   chromatin inactivation; DNA binding site; histone deacetylation.
XX
OS   Homo sapiens.
XX
PN   WO200102019-A2.
XX
PD   11-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-GB02497.
XX
PR   30-JUN-1999;   99GB-0015126.
XX
```

PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Buluwela L, Ali S;
XX
DR WPI; 2001-138069/14.
XX
PT Suppressing expression of selected gene for treating cancer, involves
PT introducing peptide comprising nucleic acid binding portion that binds
PT to or associated with selected gene and chromatin inactivation portion
PT -
XX
PS Disclosure; Fig 5; 65pp; English.
XX
CC The present invention describes a method of suppressing the expression of
CC a gene of interest by introducing into the cell a protein containing a
CC DNA binding site and a chromatin inactivation portion, or a nucleotide
CC encoding such a peptide. Preferably the chromatin inactivation portion is
CC part of the histone deacetylation (HDAC) complex. The method can be used
CC in disease treatment, for example in the treatment of cancer by the
CC suppression of oncogenes, and in the production of disease models.
XX
SQ Sequence 482 AA;

Query Match 56.7%; Score 1544; DB 22; Length 482;
Best Local Similarity 59.4%; Pred. No. 4.1e-152;
Matches 291; Conservative 67; Mismatches 98; Indels 34; Gaps 7;

```
Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | :|||||:|| :|||||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 gtrrkvcyyydgdvgnyyyyqqghpmkphrirmthnlllnyglyrkmeiyrphkanaeemt 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::||||||||||||||: |||   ||||  :|||| 
Db      66 kyhsddyikflrsirpdnmseyskqmqrfnvgedcpvfdglfefcqlstggsvasavkln 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
               |||:|||||||||| |||||||||||||||||||| |:||||:|||||||||||| 
Db     126 kqqtdiavnwagglhhakkseasgfcyvndivlailellkyhqrvlyididihhgdgvee 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| ||||||||||:|:|||||| ::||| | ||||::| || ||||||||  :|||:| 
Db     186 afyttdrvmtvsfhkygeyfpgtgdlrdigagkgkyyavnyplrdgiddesyeaifkpvm 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
            ||||:|:| |||||||:|||||||||||||:|||||:||:|::|||:|:|:|||||||| 
Db     246 skvmemfqpsavvlqcgsdslsgdrlgcfnltikghakcvefvksfnlpmlmlggggyti 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           ||||||| ||| |||  |: :::| ::|:|||||: ||::|||| |:|: : ||:|: 
Db     306 rnvarcwtyetavaldteipnelpyndyfeyfgpdfklhispsnmtnqntneylekikqr 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD--------------PDSDM 418
           |  ||  | ||| |  |  | |    | |||::| |||                 ||: 
Db     366 lfenlrmlphapgvqmqaipedaipeesgdededdpdkrisicssdkriaceeefsdsee 425

Qy     419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
           : :  ||   :     ||| |    | |:||      | |    || |   ||
Db     426 egeggrknssnfkkakrvkte----dekekd-----peekk----evteeektkeekpea 472

Qy     474 VGVEEASVKM 483
            ||:|  ||:
Db     473 kgvke-evkl 481
```

RESULT 13
AAB49954
ID AAB49954 standard; Protein; 482 AA.
XX
AC AAB49954;
XX
DT 08-MAR-2001 (first entry)
XX
DE Human histone deacetylase HDAC-1.
XX
KW Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
KW HDAC-D; cell cycle; tumourigenesis; cancer; inhibitor; antisense;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200071703-A2.
XX
PD 30-NOV-2000.
XX
PF 03-MAY-2000; 2000WO-IB01252.
XX
PR 03-MAY-1999; 99US-0132287.
XX
PA (METH-) METHYLGENE INC.
XX
PI Macleod AR, Li Z, Besterman JM;
XX
DR WPI; 2001-016407/02.
DR N-PSDB; AAC89554.
XX
PT Antisense oligonucleotide that inhibits expression of a histone
PT deacetylase, useful for treating and/or alleviating the symptoms of
PT neoplasia, or for inhibiting neoplastic cell growth in an animal -
XX
PS Disclosure; Page 49-50; 125pp; English.
XX
CC The present invention provides inhibitors of histone deacetylase enzymes
CC such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These
CC inhibitors may be antisense strands or they may be compounds identified
CC by contacting the enzyme with the compound and measuring the resulting
CC enzyme activity. These inhibitors are useful for treating cancers and for
CC identifying which histone deacetylase is involved in a neoplasia.
XX
SQ Sequence 482 AA;

Query Match 56.7%; Score 1544; DB 22; Length 482;
Best Local Similarity 59.4%; Pred. No. 4.1e-152;
Matches 291; Conservative 67; Mismatches 98; Indels 34; Gaps 7;

```
Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | :|||||:|| :|||||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 gtrrkvcyyydgdvgnyyyyqqghpmkphrirmthnlllnyglyrkmeiyrphkanaeemt 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: ||||| |:   :  :|::||||||||||||||: |||   ||||  :|||| 
Db      66 kyhsddyikflrsirpdnmseyskqmqrfnvgedcpvfdglfefcqlstggsvasavkln 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
               |||:|||||||||| |||||||||||||||||||| |:||||:|||||||||||| 
Db     126 kqqtdiavnwagglhhakkseasgfcyvndivlailellkyhqrvlyididihhgdgvee 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| ||||||||||:|:|||||| ::||| | ||||::| || ||||||||  :|||:| 
Db     186 afyttdrvmtvsfhkygeyfpgtgdlrdigagkgkyyavnyplrdgiddesyeaifkpvm 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
            ||||:|:| |||||||:|||||||||||||:|||||:||:|::|||:|:|:|||||||| 
Db     246 skvmemfqpsavvlqcgsdslsgdrlgcfnltikghakcvefvksfnlpmlmlggggyti 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           ||||||| ||| |||  |: :::| ::|:|||||: ||::|||| |:|: : ||:|: 
Db     306 rnvarcwtyetavaldteipnelpyndyfeyfgpdfklhispsnmtnqntneylekikqr 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD--------------PDSDM 418
           |  ||  | ||| |  |  | |    | |||::| |||                 ||: 
Db     366 lfenlrmlphapgvqmqaipedaipeesgdededdpdkrisicssdkriaceeefsdsee 425

Qy     419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
           : :  ||   :     ||| |    | |:||      | |    || |   ||
```

```
Db    426 egeggrknssnfkkakrvkte----dekekd-----peekk----evteeektkeekpea 472

Qy    474 VGVEEASVKM 483
          ||:|  ||:
Db    473 kgvke-evkl 481


RESULT  14
AAB56985
ID   AAB56985 standard; Protein; 488 AA.
XX
AC   AAB56985;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human prostate cancer antigen protein sequence SEQ ID NO:1563.
XX
KW   Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW   neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
KW   vulnerary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
KW   antibacterial; gene therapy; neural; immune; reproductive; renal;
KW   gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW   wound; infectious disease.
XX
OS   Homo sapiens.
XX
PN   WO200055174-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05988.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587513/55.
DR   N-PSDB; AAF16188.
XX
PT   Prostate cancer associated gene sequences, referred to as prostate
PT   cancer antigens, useful for treatment, prevention, and diagnosis of
PT   disorders such as prostate cancer -
XX
PS   Claim 11; Page 2005-2006; 2338pp; English.
XX
CC   AAF15566 to AAF16505 encode the human prostate cancer associated
CC   proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC   The prostate cancer antigens can have neuroprotective, cytostatic,
CC   cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal,
CC   nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC   and can be used in gene therapy. The prostate cancer antigen
CC   polynucleotides may be used for detection of prostate cancer, chromosome
CC   identification, as chromosome markers, and for numerous other diagnostic
CC   or research purposes. The prostate cancer antigens may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC   AAB57303 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   488 AA;


  Query Match           56.7%;  Score 1544;  DB 21;  Length 488;
  Best Local Similarity  59.4%;  Pred. No. 4.2e-152;
  Matches  291;  Conservative   67;  Mismatches   98;  Indels   34;  Gaps    7;

Qy     14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
          | :|||||:|| :|||||||||||||||||||||| || :||| : |:: :|  |   ::
Db     12 gtrrkvcyyydgdvgnyyygqghpmkphrirmthnlllnyglyrkmeiyrphkanaeemt 71

Qy     74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
          ::|:|||: ||||| |:   :  :|::|||||||||||||: |||   ||||  :||||
Db     72 kyhsddyikflrsirpdnmseyskqmqrfnvgedcpvfdglfefcqlstggsvasavkln 131

Qy    134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
           |||:|||||||||| ||||||||||||||||||||| |:||||:|||||||||||||
Db    132 kqqtdiavnwagglhhakkseasgfcyvndivlailellkyhqrvlyididihhgdgvee 191

Qy    194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
          ||| ||||||||||:|:|||||| ::||| | ||||::| || ||||||||  :|||:|
Db    192 afyttdrvmtvsfhkygeyfpgtgdirdigagkgkyyavnyplrdgiddesyeaifkpvm 251

Qy    254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
           ||||:|:| |||||||:|||||||||||||:|||||:||:|::|||:|:|:||||||||
Db    252 skvmemfqpsavvlqcgsdslsgdrlgcfnltikghakcvefvksfnlpmlmlggggyti 311

Qy    314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
          |||||| ||| |||  |: :::| ::|:|||||: ||::|||| |:|: : ||:|:
Db    312 rnvarcwtyetavaldteipnelpyndyfeyfgpdfklhispsnmtnqntneylekikqr 371

Qy    374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD--------------PDSDM 418
          |  ||  | ||| |  |  | |    |  |||::| |||                 ||:
Db    372 lfenlrmlphapgvqmqaipedaipeesgdededdpdkrisicssdkriaceeefsdsee 431

Qy    419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
          : :  ||   :     ||| |    | |:||      |  |    || |   ||
Db    432 egeggrknssnfkkakrvkte----dekekd-----peekk----evteeektkeekpea 478

Qy    474 VGVEEASVKM 483
           ||:|  ||:
Db    479 kgvke-evkl 487


RESULT  15
AAR88919
ID   AAR88919 standard; Protein; 482 AA.
XX
AC   AAR88919;
XX
DT   09-SEP-1996  (first entry)
XX
DE   RPDL, a human foetal lung-derived transcriptional control protein.
XX
KW   Transcription; expression; control; diagnosis; study; cancer;
KW   mammary; gastric carcinoma; chromosome 1p34.1.
XX
OS   Homo sapiens.
XX
PN   EP708112-A1.
XX
PD   24-APR-1996.
XX
PF   21-SEP-1995;   95EP-0114884.
XX
PR   22-SEP-1994;   94JP-0227876.
XX
PA   (CANC-) CANCER INST.
PA   (EISA ) EISAI CO LTD.
XX
PI   Furukawa Y,  Nakamura Y;
XX
DR   WPI; 1996-202003/21.
DR   N-PSDB; AAT12940.
XX
PT   New human transcriptional control protein RPDL - used to develop
PT   prods. for study and diagnosis involving the protein, partic. for
PT   gene analysis
XX
PS   Claim 1; Page 10-12; 15pp; English.
XX
CC   AAR88919 is a transcriptional control protein, RPDL, derived from a
CC   human foetal lung cDNA library. The gene encoding RPDL is localised
```

```
CC   at 1p34.1 on the short arm of chromosome 1, this is a region where a
CC   deletion is recognised in mammary and gastric carcinomas. The RPDL
CC   gene is an important gene and has been found to be expressed in all
CC   the following human tissues: heart, kidney, liver, lung, pancreas,
CC   placenta, skeletal muscle, large intestine, peripheral leukocyte,
CC   ovary, prostate, small intestine, spleen, testis and thymus gland;
CC   but is not present in brain tissue. RPDL cDNA exhibits homology with
CC   the yeast transcription factor RPD3 and is believed to have a similar
CC   function. RPDL and DNA encoding it can be used in studying the
CC   interactions of the protein with DNA and other proteins in the cell.
XX
SQ   Sequence   482 AA;


  Query Match            56.2%;  Score 1530;  DB 17;  Length 482;
  Best Local Similarity  59.2%;  Pred. No. 1.2e-150;
  Matches  290;  Conservative   67;  Mismatches   99;  Indels   34;  Gaps    7;

Qy      14 GVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQVLKPFPARERDLC 73
           | :|||||:|| :|||||||||||||||||||| || :||| : |:: :|  |   ::
Db       6 gtrrkvcyyydgdvgnyyygqghpmkphrirmthnlllnyglyrkmeiyrphkanaeemt 65

Qy      74 RFHADDYVSFLRSITPETQQDQIRQLKRFNVGEDCPVFDGLYSFCQTYAGGSVGGSVKLN 133
           ::|:|||: |||| |:   :  :|::|||||||||||||: |||   |||| :|||| 
Db      66 kyhsddyikflrsirpdnmseyskqmqrfnvgedcpvfdglfefcqlstggsvasavkln 125

Qy     134 HGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVLYVDIDIHHGDGVEE 193
               |||:|||||||||| ||||||||||||||||||| |:||||:|||||||||| 
Db     126 kqqtdiavnwagglhhakkseasgfcyvndivlailellkyhqrvlyididihhgdgvee 185

Qy     194 AFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDGIDDESYHLLFKPIM 253
           ||| |||||||||||:|:|||||| ::||| | |||||::| || |||||||  :|||:| 
Db     186 afyttdrvmtvsfhkygeyfpgtgdlrdigagkgkyyavnyplrdgiddesyeaifkpvm 245

Qy     254 GKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSFNVPLLLLGGGGYTI 313
            ||||:|:| |||||||:|||||||||||||:|||||:||:|::|||:|:|:|||||||| 
Db     246 skvmemfqpsavvlqcgsdslsgdrlgcfnltikghakcvefvksfnlpmlmlggggyti 305

Qy     314 RNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNMENKNSRQMLEEIRND 373
           |||||| ||| |||  |: :::| ::|:|||||||: ||::|||| |:|: : ||:|: 
Db     306 rnvarcrtyetavaldteipnelpyndyfeyfgpdfklhispsnmtnqntneylekikqr 365

Qy     374 LLHNLSKLQHAPSVPFQERPPDTETPEV-DEDQEDGDKRWD----------PDSDM 418
           |  ||  | ||| |  |  | |    |  |||::| |||              ||: 
Db     366 lfenlrmlphapgvqmqaipedaipeesgdededdpdkrisicssdkriaceeefsdsee 425

Qy     419 DVDDDRKPIPS-----RVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVNP 473
           : :  ||   :     ||| |    | |:||      |  |    || |   ||
Db     426 egeggrknssnfkkakrvkte----dekekd-----peekk----evteeektkeekpea 472

Qy     474 VGVEEASVKM 483
           ||:|  ||:
Db     473 kgvke-evkl 481


Search completed: April 26, 2002, 17:52:13
Job time: 81 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:57:27 ; Search time 44.51 Seconds
(without alignments)
1547.838 Million cell updates/sec

Title: US-09-645-337-4
Perfect score: 2544
Sequence: 1 MEADESGISLPSGPDGRKRR.....TGEDEMDDDNPEPDVNPPSS 471

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2542	99.9	471	10	Q9FVE5	Q9fve5 arabidopsis
2	2535	99.6	471	10	Q9FML2	Q9fml2 arabidopsis
3	1738	68.3	458	10	Q9ZTP8	Q9ztp8 zea mays (m
4	1560.5	61.3	501	10	Q9SZL3	Q9szl3 arabidopsis
5	1522	59.8	500	10	Q9LKG1	Q9lkg1 mesembryant
6	1496.5	58.8	493	10	Q9AXF0	Q9axf0 oryza sativ
7	1476	58.0	521	5	O77213	O77213 drosophila
8	1476	58.0	521	5	Q9VZA1	Q9vza1 drosophila
9	1373	54.0	687	3	Q9P4F5	Q9p4f5 emericella
10	1369	53.8	409	10	Q9FH09	Q9fh09 arabidopsis
11	1305.5	51.3	465	5	O62339	O62339 caenorhabdi
12	1292.5	50.8	428	11	Q99PA0	Q99pa0 rattus norv
13	1288	50.6	444	5	Q9GUA8	Q9gua8 cryptospori
14	1284	50.5	648	3	Q9C1C6	Q9c1c6 cochliobolu
15	1281	50.4	449	5	Q9XYC7	Q9xyc7 plasmodium
16	1277.5	50.2	428	11	Q9JM08	Q9jm08 mus musculu
17	1275.5	50.1	428	11	Q9JLX5	Q9jlx5 mus musculu
18	1275	50.1	405	3	O59702	O59702 schizosacch
19	1233.5	48.5	438	5	Q9VNC2	Q9vnc2 drosophila
20	1229	48.3	566	3	Q9HDT2	Q9hdt2 ustilago ma
21	1207	47.4	419	10	Q9M1N6	Q9m1n6 arabidopsis
22	1151.5	45.3	437	5	Q9GU59	Q9gu59 cryptospori
23	1072	42.1	429	5	Q9GT74	Q9gt74 tetrahymena
24	994.5	39.1	481	3	Q9P4F4	Q9p4f4 emericella
25	820.5	32.3	377	11	Q9D0K6	Q9d0k6 mus musculu
26	813	32.0	377	4	Q9NP76	Q9np76 homo sapien
27	808	31.8	377	4	Q9NYH4	Q9nyh4 homo sapien
28	748.5	29.4	428	5	Q9GRP1	Q9grp1 leishmania
29	596.5	23.4	256	11	Q99PA2	Q99pa2 rattus norv
30	586	23.0	223	4	Q9H368	Q9h368 homo sapien
31	457.5	18.0	375	2	O67135	O67135 aquifex aeo
32	424.5	16.7	367	2	Q9WX04	Q9wx04 streptomyce
33	410.5	16.1	389	2	Q9K7X1	Q9k7x1 bacillus ha
34	371.5	14.6	389	2	Q99TC9	Q99tc9 staphylococ
35	319.5	12.6	380	2	Q9HXM1	Q9hxm1 pseudomonas
36	305	12.0	310	2	O67877	O67877 aquifex aeo
37	299.5	11.8	158	10	Q9LXN8	Q9lxn8 arabidopsis
38	290.5	11.4	883	5	Q9XYX1	Q9xyx1 drosophila
39	277	10.9	1108	10	Q9FNQ7	Q9fnq7 arabidopsis
40	273.5	10.8	1114	11	Q9JL73	Q9jl73 mus musculu
41	269.5	10.6	425	10	Q9SJE6	Q9sje6 arabidopsis
42	269.5	10.6	796	5	O17323	O17323 caenorhabdi
43	268	10.5	338	1	Q9HSP7	Q9hsp7 halobacteri
44	266.5	10.5	577	10	Q9LS38	Q9ls38 arabidopsis
45	265	10.4	878	4	Q9UFU7	Q9ufu7 homo sapien

ALIGNMENTS

```
RESULT   1
Q9FVE5
ID   Q9FVE5      PRELIMINARY;      PRT;   471 AA.
AC   Q9FVE5;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE.
GN   RPD3B.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RA   Wu K., Malik K., Tian L., Brown D., Miki B.;
RT   "Functional analysis of RPD3 histone deacetylase homologs in
RT   Arabidopsis thaliana.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF195548; AAG28475.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   471 AA;  52720 MW;  371BF7040E508849 CRC64;
```

Query Match 99.9%; Score 2542; DB 10; Length 471;
Best Local Similarity 100.0%; Pred. No. 5.3e-198;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60

Qy      61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120

Qy     121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180

Qy    181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240

Qy    241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300

Qy    301 RSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 RSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360

Qy    361 PMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 PMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWS 420

Qy    421 GTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
```

RESULT 2
Q9FML2
ID Q9FML2 PRELIMINARY; PRT; 471 AA.
AC Q9FML2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HISTONE DEACETYLASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones.";
RL DNA Res. 4:401-414(1997).
DR EMBL; AB008265; BAB10553.1; -.
DR InterPro; IPR000286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 471 AA; 52651 MW; CA16C2640D1B1732 CRC64;

Query Match 99.6%; Score 2535; DB 10; Length 471;
Best Local Similarity 99.6%; Pred. No. 1.9e-197;
Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60

Qy     61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120

Qy    121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180

Qy    181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240

Qy    241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300

Qy    301 RSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 RSYNVPLMVLGGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360

Qy    361 PMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 PMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRIWS 420

Qy    421 GTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
```

RESULT 3
Q9ZTP8
ID Q9ZTP8 PRELIMINARY; PRT; 458 AA.
AC Q9ZTP8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HISTONE DEACETYLASE.
GN HD1B.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CUZCO;
RA Pipal A., Wegener S.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF045473; AAD10139.1; -.
DR Mendel; 39025; Zeama;3043;39025.
DR InterPro; IPR000286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 458 AA; 50940 MW; A62775068225BE79 CRC64;

Query Match 68.3%; Score 1738; DB 10; Length 458;
Best Local Similarity 69.3%; Pred. No. 8.5e-133;
Matches 328; Conservative 51; Mismatches 72; Indels 22; Gaps 7;

```
Qy      1 MEADESGISLPS---GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHL 57
          | |   |:||||   | |  :|||||||||:|||||||||||||||||||||::|| |
Db      1 MAASGEGVSLPSPAGGEDAHRRRVSYFYEPSIGDYYYGQGHPMKPHRIRMAHSLVVHYGL 60

Qy     58 HRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSA--ARNLRRFNVGEDCPVF 115
          || ||:|||  |  :|| |||| :|| |||| :    |:|     | ::|||||||||||
Db     61 HRLLELSRPYPASEADIRRFHSDDYVAFLASAT----GNPGVLDPRAIKRFNVGEDCPVF 116

Qy    116 DGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILEL 175
          |||| ||:|||||||||||||||| |||| :|| |||||||||||||||||||||| ||||
Db    117 DGLFPFCQASAGGSIGAAVKLNRGDADITVNWAGGLHHAKKSEASGFCYVNDIVLAILEL 176

Qy    176 LKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYA 235
          ||  :||||:|||||||||||||:||:|||||||||:|||||||||| |||| :||:||
Db    177 LKFHRRVLYVDIDVHHGDGVEEAFFTTNRVMTVSFHKYGDFFPGTGHITDVGAAEGKHYA 236

Qy    236 LNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHAD 295
          |||||:||:|| :|| ||: :|:|||||||| ||||||||||:|||||||||||||||
Db    237 LNVPLSDGIDDTTFRGLFQCIIKKVMEVYQPDVVVLQCGADSLAGDRLGCFNLSVKGHAD 296

Qy    296 CLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTL 355
          ||||||||||:|||| ||||||||||||||||||||||||||||:|:|||||||||
Db    297 CLRFLRSYNVPMMVLGGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNDYYEYFGPDYTL 356

Qy    356 HVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRP- 414
```

```
          ‡: |  :|||||  ||:|  |:|  :||  ‡‡  : |  ||  ||    |       :| |:||: ||
Db    357 HIQPKSVENLNTTKDLENIKNMILENLSKIEHVPSTQFHDRPSDPEAPEEKEEDMDKRPP 416

Qy    415 -KPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDV 466
           : ‡ ||| | |:||::| |            :  :| |   :| ‡: |: |:
Db    417 QRSRLWSGGA-YDSDTEDPDSL---------KSEGKDVTANFQMKDE-PKDDL 458


RESULT   4
Q9SZL3
ID   Q9SZL3      PRELIMINARY;      PRT;   501 AA.
AC   Q9SZL3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
GN   F20D10.250 OR AT4G38130.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Wedler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,
RA   Mayer K.F.X., Schueller C.;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Wedler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,
RA   Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035538; CAB37553.1; -.
DR   EMBL; AL161593; CAB80478.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   501 AA;  56023 MW;  857D2E3D16B7CC1F CRC64;


  Query Match             61.3%;  Score 1560.5;  DB 10;  Length 501;
  Best Local Similarity   60.7%;  Pred. No. 2.4e-118;
  Matches  292;  Conservative   64;  Mismatches   94;  Indels   31;  Gaps    5;

Qy      4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
          |  | || ||||| ||:| |‡|:| :|:|||||||||||||||||| |:|: ‡‡ | : :::
Db      2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61

Qy     64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
           :|  |   |: |‡|: :|| || |::||:  ‡    | |:|||||||||||||: ||:
Db     62 LKPFPARDRDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119

Qy    124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
            ||||:| :|‡‡    |||||| ‡‡‡‡‡‡‡‡ |||||||||||||| ||||||  :|‡‡
Db    120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179

Qy    184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
          |:|||:|||||||||||| |||‡‡‡‡‡‡‡‡‡‡‡‡‡:‡‡‡‡‡‡‡:|:|   ||||:‡|||||:||
Db    180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239

Qy    244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
          :|‡‡‡:  ||:|:: |‡‡|:::| |||||||||||||||||||||||:‡‡‡‡:‡::|:||:
Db    240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299

Qy    304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363
```

```
          ||||::||| |||||||||||||| ||:||| ::|:| :||:|||||||||| || ||
Db    300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359

Qy    364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXWSG 421
          | |: : :| ||| ||  || | ‡‡‡‡ ||  ||      :|| ::| : |  |
Db    360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDS--- 416

Qy    422 TATYESDSDDDDKP------------------LHGYSCRG-GATTDRDSTGEDEMDDDN 461
            : | ||| ||                  ‡ |   || |   : | :|  ::   |
Db    417 ----DMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472

Qy    462 P 462
          |
Db    473 P 473


RESULT   5
Q9LKG1
ID   Q9LKG1      PRELIMINARY;      PRT;   500 AA.
AC   Q9LKG1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
GN   HDEAC1.
OS   Mesembryanthemum crystallinum (Common ice plant).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX   NCBI_TaxID=3544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Scharte J., Baur B.;
RT   "Molecular cloning of histone deacetylase from Mesembryanthemum
RT   crystallinum.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF282858; AAF82385.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   500 AA;  56264 MW;  CFD4FE0525209ABC CRC64;


  Query Match             59.8%;  Score 1522;  DB 10;  Length 500;
  Best Local Similarity   61.5%;  Pred. No. 3.2e-115;
  Matches  288;  Conservative   67;  Mismatches   95;  Indels   18;  Gaps    5;

Qy      4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
          ‡  ‡ |||||||| ||:|||||:| :|:|||||||||||||||| |:|: || | : :::
Db      2 DTGGNSLPSGPDGVKRKVSYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61

Qy     64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
           :|  |   |: |‡|: :|| ||  ::‡‡   ‡    ‡ ‡:|||||||||||||: ‡‡:
Db     62 LKPVPARDRDLCRFHADDYVSFLRGITPEMQQD--QLRQLKRFNVGEDCPVFDGLYSFCQ 119

Qy    124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
            ||||:| |||||    |||||| |‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡‡ ||||||: :|‡|
Db    120 TYAGGSVGGAVKLNHGLCDIAINWAGGLHHAKKSEASGFCYVNDIVLAILELLKVHERVL 179

Qy    184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
          |:|||:|||||||||||:|||||||||||:|||    :|  |||||:|||||:||
Db    180 YVDIDIHHGDGVEEAFYTTERVMTVSFHKFGDYFPGTETFVYIGYGKGKYYSLNVPLDDG 239

Qy    244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
          :|‡‡‡‡::|||:|:: |‡‡|:::| ||||||||||||||||||||||||:‡‡‡‡:|:‡‡‡ ‡:
Db    240 IDDESYQSLFKPIMGKVMEIFKPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKYMISF 299

Qy    304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363
          |||‡:::|| ‡‡‡‡ |‡|‡‡‡‡‡‡ ||:| | |:|:| :||:|||||||||| || ||
Db    300 NVPLLLVGGGGYTIRYVARCWCYETGVALGFELDDKIPQHEYYEYFGPDYTLHVAPSNME 359

Qy    364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPED--DMETRPKPRXWSG 421
```

```
          | |:   :: ||  ||| || | |||||||  ||   : :| ||  | | |  |
Db    360 NKNSRPMLDDIRGKLLEYLSSLQHAPSVQFQERPPEAEIPEEDEDQHDPEGRWDP----- 414

Qy    422 TATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPP 469
              :|| : ||: :   | | :   |:    |  | |: :  |: |
Db    415 ----DSDMEVDDRSI---SIR--SRVKREIMEPDLKDRDDVKDAVDHP 453


RESULT   6
Q9AXF0
ID   Q9AXF0      PRELIMINARY;      PRT;   493 AA.
AC   Q9AXF0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE HD1.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. YUANFENGZAO;
RA   Song F., Goodman R.M.;
RT   "Molecular characterization of a rice histone deacetylase gene
RT   OsHD1.";
RL   Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF332875; AAK01712.1; -.
SQ   SEQUENCE   493 AA;  55054 MW;  0EE300BB42243C13 CRC64;


  Query Match             58.8%;  Score 1496.5;  DB 10;  Length 493;
  Best Local Similarity   60.5%;  Pred. No. 3.7e-113;
  Matches  294;  Conservative   64;  Mismatches   97;  Indels   31;  Gaps   11;

Qy      3 ADESGISLP-SGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
            |   | ||| :| || |||| :||:  :|:|| ||||||||||||| |:|: || |  ::
Db      5 AGGGGNSLPTAGADGAKRRVCHFYDAEVGNYYCGQGHPMKPHRIRMTHALLAHYGLLDQM 64

Qy     62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDF 121
          :: :|  |   |: ||| :|| || ||:||:  |    | |:|||||||||||||: |
Db     65 QVLKPHPARDRDLCRFHADDYVAFLRSVTPETQQD--QIRALKRFNVGEDCPVFDGLYSF 122

Qy    122 CRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKR 181
          |:  ||||:| |||||    |||||| |||||||| |||||||||||||| ||||||  :|
Db    123 CQTYAGGSVGGAVKLN-HGHDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKYHQR 181

Qy    182 VLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLN 241
          |||:|||:|||||||||||||||||||||||||:||||| |||:|  :|||| |||||:
Db    182 VLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGYSEGKYYCLNVPLD 241

Qy    242 DGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLR 301
          ||:|||:|::|:|:|:| ||||:|:| |||||||||||||||||||| ||||:|::|:|
Db    242 DGIDDDSYQSIFKPIISKVMEMYRPGAVVLQCGADSLSGDRLGCFNLSGKGHAECVKFMR 301

Qy    302 SYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP 361
          |:|||::||| || |||||||||| ||:| |  :|:| |||||||||:|:| |  |
Db    302 SFNVPLLLLGGGGYAIRNVARCWCYETGVALGHELTDKMPPNEYFEYFGPEYSLFVAASN 361

Qy    362 MENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSG 421
          ||| || | :| |:  :|: || | |||||||:   |  : | ||::| |  |  |
Db    362 MENRNTNKQLEEIKCNILDNLSKLQHAPSVQFEERIPETK-LPEPDEDQED-PDER---- 415

Qy    422 TATYESDSD---DDDKPLHGYSCRG---GATTDRDSTGEDEMDDDN----------PEPD 465
             :: |||   || ||: |:| |         |: | | |  | :          |||
Db    416 ---HDPDSDMVLDDHKPM-GHSARSLIRNIEVKREIT-ESEAKDQHGKRLTTEHKGPEPM 470

Qy    466 VNPPSS 471
           : | |
Db    471 ADDPGS 476
```

```
RESULT   7
O77213
ID   O77213      PRELIMINARY;      PRT;   521 AA.
AC   O77213;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PUTATIVE HISTONE DEACETYLASE.
GN   RPD3 OR CG7471.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Mottus R.C., Sobel R.E., Grigliatti T.A.;
RT   "DmHDAC1.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF086715; AAC61494.1; -.
DR   FlyBase; FBgn0015805; Rpd3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   521 AA;  58344 MW;  B0F6503D42A1BA32 CRC64;


  Query Match             58.0%;  Score 1476;  DB 5;  Length 521;
  Best Local Similarity   58.6%;  Pred. No. 1.8e-111;
  Matches  273;  Conservative   76;  Mismatches   89;  Indels   28;  Gaps    5;

Qy     18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
          |:|| |:|:  ||:|||||||||||||||| |:|:::| |:|::|| ||  | | :: :|
Db      6 KKRVCYYYDSDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKATADEMTKF 65

Qy     78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
          || ||| || |: |:::| : :  : ::|||||||||||||::||: |||||: ||||||
Db     66 HSDEYVRFLRSIRPDNMSEYN--KQMQRFNVGEDCPVFDGLYEFCQLSAGGSVAAAVKLN 123

Qy    138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKYHQRVLYIDIDVHHGDGVEE 197
          :| ::| ||||||||||||||||||||||||||||||||  :|||||||||||||||||
Db    124 KQASEICINWGGGLHHAKKSEASGFCYVNDIVLGILELLKYHQRVLYIDIDVHHGDGVEE 183

Qy    198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
          |||||||||||||||:|::||||| :||:|| ||||||:|:|| |||||::: |:| |:|
Db    184 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNIPLRDGMDDDAYESIFVPII 243

Qy    258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
          |||| :|| ||||||||||||:|||||||||:|||  |: |:: ||:| :::|| ||||
Db    244 SKVMETFQPAAVVLQCGADSLTGDRLGCFNLTVKGHGKCVEFVKKYNLPFLMVGGGGYTI 303

Qy    318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNT 377
          |||:||| |||:||: ||  |:||||:|||||||: ||: || | | || : :|:|:|
Db    304 RNVSRCWTYETSVALAVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTSEYLEKIKNR 363

Qy    378 LLEQLSGLIHAPSVQFQHTP--PVNRVLDE----------PEDDMETRPKPRXWSGTATY 425
          | | |  | ||| || |  |   :|   |:          |: | : |  |
Db    364 LFENLRMLPHAPGVQIQAIPEDAINDESDDEDKVDKDDRLPQSDKDKRIVPE------NE 417

Qy    426 ESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
          |||:|: :        ||   :|   |: :    : : : |  ||
Db    418 YSDSEDEGE--------GGRRDNRSYKGQRKRPRLDKDTNSNKASS 455


RESULT   8
Q9VZA1
ID   Q9VZA1      PRELIMINARY;      PRT;   521 AA.
AC   Q9VZA1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   RPD3 PROTEIN.
GN   RPD3 OR CG7471.
```

```
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003482; AAF47924.1; -.
DR   FlyBase; FBgn0015805; Rpd3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   521 AA;  58330 MW;  B0F6503D42A1BCE9 CRC64;
```

```
  Query Match             58.0%;  Score 1476;  DB 5;  Length 521;
  Best Local Similarity   58.6%;  Pred. No. 1.8e-111;
  Matches  273;  Conservative   76;  Mismatches   89;  Indels   28;  Gaps    5;

Qy      18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
           |:|| |:|:  ||:||||||||||||||||| |:|:::| |:|::|| ||  | | :: :|
Db       6 KKRVCYYYDSDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKATADEMTKF 65

Qy      78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
           || ||| || |: |::| : :  : ::|||||||||||||::||: |||||: ||||||
Db      66 HSDEYVRFLRSIRPDNMSEYN--KQMQRFNVGEDCPVFDGLYEFCQLSAGGSVAAAVKLN 123

Qy     138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEE 197
           :| ::| ||||||||||||||||||||||||||||||||  :|||||||||||||||||
Db     124 KQASEICINWGGGLHHAKKSEASGFCYVNDIVLGILELLKYHQRVLYIDIDVHHGDGVEE 183

Qy     198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
           |||||||||||||||:|::|||||  :||:|| |||||||:|:|| ||||||::: |:| |:|
```

```
Db     184 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNIPLRDGMDDDAYESIFVPII 243

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
           |||| :|| ||||||||||||:|||||||||:|||| |: |:: ||:| :::|| ||||
Db     244 SKVMETFQPAAVVLQCGADSLTGDRLGCFNLTVKGHGKCVEFVKKYNLPFLMVGGGGYTI 303

Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNT 377
           |||:||| |||:||: ||  |:||||:||||||||: ||: || | | || : :|:|:|
Db     304 RNVSRCWTYETSVALAVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTSEYLEKIKNR 363

Qy     378 LLEQLSGLIHAPSVQFQHTP--PVNRVLDE----------PEDDMETRPKPRXWSGTATY 425
           | | |  | ||| || | |  :|  |:          |: | : |  |
Db     364 LFENLRMLPHAPGVQIQAIPEDAINDESDDEDKVDKDDRLPQSDKDKRIVPE------NE 417

Qy     426 ESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
           |||:|: :        ||  :|  |: :   : : : |  ||
Db     418 YSDSEDEGE--------GGRRDNRSYKGQRKRPRLDKDTNSNKASS 455
```

```
RESULT   9
Q9P4F5
ID   Q9P4F5      PRELIMINARY;      PRT;   687 AA.
AC   Q9P4F5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE RPD3A.
GN   RPD3A.
OS   Emericella nidulans (Aspergillus nidulans).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=5072;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20461766; PubMed=11004483;
RA   Graessle S., Dangl M., Haas H., Mair K., Trojer P., Brandtner E.M.,
RA   Walton J.D., Loidl P., Brosch G.;
RT   "Characterization of two putative histone deacetylase genes from
RT   Aspergillus nidulans.";
RL   Biochim. Biophys. Acta 1492:120-126(2000).
DR   EMBL; AF163862; AAF80489.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   687 AA;  75431 MW;  CFC7C65352366451 CRC64;
```

```
  Query Match             54.0%;  Score 1373;  DB 3;  Length 687;
  Best Local Similarity   49.9%;  Pred. No. 6.2e-103;
  Matches  266;  Conservative   86;  Mismatches  111;  Indels   70;  Gaps   10;

Qy       6 SGISLPSGP----------DGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHY 55
           || | |:||           | :||:|||:  :|:| |  ||||||||||| ||||:|
Db       3 SGTSGPAGPPLDPIDLNVSGDRSKRVAYFYDSDVGNYAYVSGHPMKPHRIRMTHSLVMNY 62

Qy      56 HLHRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVF 115
           |:::||  |   |   :: :||: ||:|||: |:|::|   : |:   ::|||:|||||
Db      63 SLYKKMEIYRAKPASKFEMTQFHTDEYIDFLSKVTPDNM--DAFAKEQSKYNVGDDCPVF 120

Qy     116 DGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILEL 175
           ||||:||  |||||:  | :|||   |||:|| |||||||||||||||||||||||||||
Db     121 DGLFEFCGISAGGSMEGAARLNRNKCDIAVNWAGGLHHAKKSEASGFCYVNDIVLGILEL 180

Qy     176 LKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYA 235
           |:  :||||:|||||||||||||||||||||||||||:|::|||| :||:|  :|||||
Db     181 LRFKQRVLYVDIDVHHGDGVEEAFYTTDRVMTVSFHKYGEYFPGTGELRDIGVGQGKYYA 240

Qy     236 LNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHAD 295
           :| || ||:|| |::|:| |:|: ||| |:|||||||| ||||||||||||::|||:
Db     241 VNFPLRDGIDDVSYKSIFEPVIKSVMEWYRPEAVVLQCGGDSLSGDRLGCFNLSMRGHAN 300

Qy     296 CLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTL 355
           |:::::|:|:| :::|| |||:|||||  | :|| : ||    ::||||:|:||| ||| |
```

```
Db    301 CVKYVKSFNLPTLIVGGGGYTMRNVARTWAFETGILVGDNLGSELPYNDYYEYFAPDYEL 360

Qy    356 HVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPP---VNRVLDEPE---DD 409
          | || |:| || : :::||  ::| |     |||||    |    |: : || |   ||
Db    361 DVRPSNMDNANTREYLDKIRTQVVENLKRTAFAPSVQMTDVPREPLVDGMDDEAEAALDD 420

Qy    410 M-ETRPKPRXWSGTATYE--------SDSDDDD---------KPLH-------------- 437
          : |   | : ::     :         |||:|:|         || |
Db    421 LDEDENKDKRFTKRRFDQYVEKPGELSDSEDEDENAANGVTRKPAHLKRRNQANYRLDLA 480

Qy    438 --GYSCRGGATTDRDSTGEDEMDD---------DNPEPD---------VNPPS 470
            |         |  |  ::|||           ||||           |||
Db    481 DSGVESGMATPQDASSVADEEMDTGTDVKITEAPGPEPDSEAQGTSSAAEPPS 533


RESULT  10
Q9FH09
ID   Q9FH09      PRELIMINARY;      PRT;   409 AA.
AC   Q9FH09;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RX   MEDLINE=20181125; PubMed=10718197;
RA   Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT   features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:31-63(2000).
DR   EMBL; AB023031; BAB09994.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   409 AA;  46017 MW;  56A4FFAF9A0734AE CRC64;


  Query Match             53.8%;  Score 1369;  DB 10;  Length 409;
  Best Local Similarity   63.9%;  Pred. No. 6.2e-103;
  Matches  260;  Conservative   50;  Mismatches   89;  Indels    8;  Gaps    3;

Qy     15 DGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDI 74
          || |||||||||| |||||||   | || |||: |:||: |:||| :||: | ||||||
Db      6 DGGKRRVSYFYEPMIGDYYYGVNQPTKPQRIRVTHNLILSYNLHRHMEINHPDLADASDF 65

Qy     75 GRFHSPEYVDFLASVSPESMGD--PSAARNLRRFNVGEDC--PVFDGLFDFCRASAGGSI 130
             :||| ||::|| ||:||:: |  || : ||:|||| |   |||  |||:||| |||||
Db     66 EKFHSLEYINFLKSVTPETVTDPHPSVSENLKRFNVDVDWDGPVFHNLFDYCRAYAGGSI 125

Qy    131 GAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVH 190
          || |||||:|||||| ||:|| || :|||| ||||:|| |||||| |||||||:|
Db    126 SAAAKLNRQEADIAINWAGGMHHVKKDKASGFGYVNDVVLAILELLKSFKRVLYIEIGFP 185

Qy    191 HGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFR 250
          ||| |||||  |||||||||||| ||    || | | |  ||:||:|| || ||:|| | |
Db    186 HGDEVEEAFKDTDRVMTVSFHKVGD----TGDISDYGEGKGQYYSLNAPLKDGLDDFSLR 241

Qy    251 SLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVL 310
          || |:| : ||:|:|| :|||||||||:||  | ||||:||| |||:::||:|||||:|
Db    242 GLFIPVIHRAMEIYEPEVIVLQCGADSLAGDPFGTFNLSIKGHGDCLQYVRSFNVPLMIL 301

Qy    311 GGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKD 370
          || |||: |||||||||||:||| : || || |:| :|| ||| ||: |:  :|||| |
Db    302 GGGGYTLPNVARCWCYETAIAVGEQLDNDLPGNDYMKYFRPDYKLHILPTNRQNLNTRLD 361

Qy    371 MERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPR 417
          :  :| ||| ||| ::||||| || ||  ::  :  | ||| |  ||
Db    362 IITMRETLLAQLSLVMHAPSVPFQDTPSSSQATEAAEVDMEKRNDPR 408


RESULT  11
O62339
ID   O62339      PRELIMINARY;   PRT;   465 AA.
AC   O62339; O62343;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   R06C1.1 PROTEIN.
GN   R06C1.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kershaw J.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z81108; CAB03240.1; -.
DR   EMBL; Z81106; CAB03240.1; JOINED.
DR   EMBL; Z81106; CAB03224.1; -.
DR   EMBL; Z81108; CAB03224.1; JOINED.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   465 AA;  52709 MW;  325E71D5197012C5 CRC64;


  Query Match             51.3%;  Score 1305.5;  DB 5;  Length 465;
  Best Local Similarity   61.9%;  Pred. No. 1.1e-97;
  Matches  234;  Conservative   67;  Mismatches   74;  Indels    3;  Gaps    2;

Qy     18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
          | ||||:|:   |::|||||||||||||:|| ||||::| |:|:| : ||: |  |:| |:
Db      7 KSRVSYYYDGDFGNFYYGQGHPMKPHRVRMTHSLIVNYGLYRKLNVMRPARASFSEITRY 66

Qy     78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
          || :|::|| :|  ::|   :    : ||:||||||||:::||: | ||| || :||
Db     67 HSDDYINFLRNVKSDNMS--TFTDQMARFSVGEDCPVFDGMYEFCQLSCGGSLAAAARLN 124

Qy    138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEE 197
          ||:::|||| |||||||||||||||| ||||| ||||||  ||||||||||||||||||
Db    125 RQESEIAINWMGGLHHAKKSEASGFCYSNDIVLAILELLKHHKRVLYIDIDVHHGDGVEE 184

Qy    198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
          |||||||||||||| |::|||| ::||||  |||||||||| ||:|| ::  :|| ::
Db    185 AFYTTDRVMTVSFHKHGEYFPGTGDLKDVGAGSGKYYALNVPLRDGVDDVTYERIFRTIM 244

Qy    258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
          :||  :|||||||||||||:||||| |||:  ||  |: :::|:||||:::|| ||||
Db    245 GEVMARFQPEAVVLQCGADSLAGDRLGVFNLTTYGHGKCVEYMKSFNVPLLLVGGGGYTI 304

Qy    318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDP-SPMENLNTPKDMERIRN 376
```

```
          |||:||| ||||:|:  |  : || ::||:|| ||| ||: | : : | |||: :::
Db    305 RNVSRCWLYETAIALNQEVSDDLPLHDYFDYFIPDYKLHIKPLAALSNFNTPEFIDQTIV 364

Qy    377 TLLEQLSGLIHAPSVQFQ 394
          ||| |  | | ||||  |
Db    365 ALLENLKQLPHVPSVQMQ 382


RESULT  12
Q99PA0
ID   Q99PA0      PRELIMINARY;      PRT;   428 AA.
AC   Q99PA0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE 3.
GN   HDAC3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=TESTIS;
RA   Wilquet V., Chavez M., Korbers R., Geerts A.;
RT   "Expression pattern of rat histone deacetylases.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF321131; AAK11184.1; -.
SQ   SEQUENCE   428 AA;  48829 MW;  7CB448CFE33041FE CRC64;


  Query Match             50.8%;  Score 1292.5;  DB 11;  Length 428;
  Best Local Similarity   55.8%;  Pred. No. 1.1e-96;
  Matches  235;  Conservative   74;  Mismatches  103;  Indels    9;  Gaps    5;

Qy     19 RRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFH 78
          : |:|||:| :|:::|| ||||||||: : |||::|| |:::: : :|  |   |: |||
Db      3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy     79 SPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNR 138
          | :|:|||  ||| :|      ::|  ||||:||||| |||:||    | |:  | :||
Db     63 SEDYIDFLQRVSPTNM--QGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNN 120

Qy    139 QDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEA 198
          :  |||||| |||||||| ||||||||||||:|||||||   ||||||||:||||||:||
Db    121 KICDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEA 180

Qy    199 FYTTDRVMTVSFHKFGD-FFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
          || |||||||||||:|: |||||| : :|||| |:|| ||||| ||:||:|:: ||:|:|
Db    181 FYLTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVI 240

Qy    258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
          :|:: |||  :|||||||||  ||||||||||::|| :|: :::|:|:||:|||| |||:
Db    241 SQVVDFYQPXCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTV 300

Qy    318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP-MENLNTPKDMERIRN 376
          ||||||| |||:: |      :|||:|||||| ||:||| | |  :|| |: : :::||
Db    301 RNVARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQ 360

Qy    377 TLLEQLSGLIHAPSVQFQHTPPVNRVLDEPED-DMETRPKPRXWSGTAT----YESDSDD 431
          |: | |  | ||||||    |      |  :: | | |     :|        |: | |:
Db    361 TIFENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDN 420

Qy    432 D 432
          |
Db    421 D 421


RESULT  13
Q9GUA8
ID   Q9GUA8      PRELIMINARY;      PRT;   444 AA.
AC   Q9GUA8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
GN   HD1.
OS   Cryptosporidium parvum.
OC   Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC   Cryptosporidiidae; Cryptosporidium.
OX   NCBI_TaxID=5807;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KSU-1;
RA   Kasper J.K., Woods K.W., Khramtsov N.V., Upton S.J.;
RT   "Cloning and analysis of a Cryptosporidium parvum gene encoding a
RT   protein with homology to probable histone deacetylase 1-2 (HD1)(RPD3
RT   homolog).";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF139991; AAG21919.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
SQ   SEQUENCE   444 AA;  50476 MW;  B1686DE715D2C02B CRC64;


  Query Match             50.6%;  Score 1288;  DB 5;  Length 444;
  Best Local Similarity   54.9%;  Pred. No. 2.6e-96;
  Matches  237;  Conservative   80;  Mismatches   95;  Indels   20;  Gaps    7;

Qy     19 RRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFH 78
          :|||||:  || |||| |||||| ||||||:||: | |:: :|| :|  :  |::  ||
Db      3 KRVSYFYDGDIGSYYYGPGHPMKPQRIRMAHNLILSYDLYKHMEIYKPHKSPQSELVYFH 62

Qy     79 SPEYVDFLASVSPESMGDPSAARNLRRFNVGE--DCPVFDGLFDFCRASAGGSIGAAVKL 136
            :|::||:|::|::  |      |:|||:||  |||||||||:| : |||||  | ||
Db     63 EEDYINFLSSINPDNSKD--FGLQLKRFNLGETTDCPVFDGLFEFQQICAGGSIDGAYKL 120

Qy    137 NRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVE 196
          | : :|| ||| |||||||:|||||||:|||||||||||||   ||:|||||||||||||
Db    121 NNEQSDICINWSGGLHHAKRSEASGFCYINDIVLGILELLKYHARVMYIDIDVHHGDGVE 180

Qy    197 EAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPL 256
          |||| : ||:|||||||:|||||| | |:|  :||||::|||||||:|||:|| |||:|:
Db    181 EAFYLSHRVLTVSFHKFGEFFPGTGDITDIGVAQGKYYSVNVPLNDGIDDDSFLSLFKPI 240

Qy    257 IQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYT 316
          | | :|||:| |:||||||||: ||||| ||||:||||:|: | : :|:||::||| |||
Db    241 ISKCIEVYRPGAIVLQCGADSVRGDRLGRFNLSIKGHAECVEFCKIFNIPLLILGGGGYT 300

Qy    317 IRNVARCWCYETAVAV---GVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMER 373
          |||||| | ||||  :    :  || :|  :|:::|| ||: ||: |  : |:|:|: :|:
Db    301 IRNVARTWAYETATILDRTDLISDN-IPLYDYYDYFAPDFKLHIPPLNLPNMNSPEHLEK 359

Qy    374 IRNTLLEQLSGLIHAPSVQFQHTPPVNRVLD------EPEDDMETRPKPRXWSG----TA 423
          |:  ::: |  | ||| |:| : |  :    |      : ::| |   :   | |
Db    360 IKAKVIDNLRYLEHAPGVEFAYVP--SDFFDREASNLQKQEDEEREEELSSWQGGGRAAG 417

Qy    424 TYESDSDDDDKP 435
          : ||  : ::||
Db    418 STESQGNHNEKP 429


RESULT  14
Q9C1C6
ID   Q9C1C6      PRELIMINARY;      PRT;   648 AA.
AC   Q9C1C6;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE 2.
GN   HDC2.
OS   Cochliobolus carbonum (Bipolaris zeicola).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
OC   Pleosporales; Pleosporaceae; Cochliobolus.
OX   NCBI_TaxID=5017;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wegener S., Walton J.D.;
RT   "HDC2, a histone deacetylase gene related to RPD3, from the
RT   filamentous fungus Cochliobolus carbonum.";
RL   Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF349677; AAK35180.1; -.
SQ   SEQUENCE   648 AA;  70940 MW;  242147C293D4EAD3 CRC64;


  Query Match             50.5%;  Score 1284;  DB 3;  Length 648;
  Best Local Similarity   50.1%;  Pred. No. 9.5e-96;
  Matches  246;  Conservative   86;  Mismatches  123;  Indels   36;  Gaps    8;

Qy       1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60
           |       : : :|   : ::|:|||:  :|:| |  ||||||||||||||||::| |: :
Db       7 MAGPTGAVHINNGTASQPKKVAYFYDSDVGNYAYVAGHPMKPHRIRMAHSLIMNYGLYTK 66

Qy      61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120
           :|| |   |   :: :||: ||::||  |:|::|     |   ::|||:||||||||:
Db      67 MEIYRAKPASKYEMTQFHTDEYIEFLHKVTPDNM--EQFTREQSKYNVGDDCPVFDGLFE 124

Qy     121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180
           ||  ||||::  | :|||   |:|:|| ||||||||||||||||||||||:|||:  :
Db     125 FCGISAGGTMEGAARLNRGKCDVAVNWAGGLHHAKKSEASGFCYVNDIVLGIIELLRYKQ 184

Qy     181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240
           ||||||||||||||||||||||||||||||:|::|||||  :|||:|   || |::| ||
Db     185 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKYGEYFPGTGELRDIGVGSGKNYSVNFPL 244

Qy     241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300
           ||: ||::|:::| |:|: ||  | |||:|||| ||||||||||||: |||:|:|::
Db     245 RDGITDETYRNIFEPVIEAVMTYYGPEAIVLQCGGDSLSGDRLGCFNLSMDGHANCVRYV 304

Qy     301 RSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360
           :|: ||::|||| |||:|||| | |||   |  :   :||:|:|:||| ||| | | ||
Db     305 KSFGVPVIVLGGGGYTMRNVARTWAYETGELVSQKMSKQLPFNDYYEYFAPDYELDVRPS 364

Qy     361 PMENLNTPKDMERIRNTLLEQL--SGLIHAPSVQFQHTPP-----VNRVLD----EPEDD 409
           ||| |:   : :|:: ::| :  :|    |||:   | |     : | :|    : |||
Db     365 NMENANSHDYLHKIKSAVIENIRRTG---RPSVEAFTTIPDVPTALGRAMDSDAEDEEDD 421

Qy     410 MET-----------RPKPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEM 457
           |:            | |     |   |::  |:| |   |   : |    |: |
Db     422 MDADENRDVRMTQRQRDKQIEHDG-ELYDASDDEDYKNNLGVRAQPGVKKRRNIT----- 475

Qy     458 DDDNPEPDVNP 468
             | | |:  |
Db     476 --DFPNPNAAP 484


RESULT  15
Q9XYC7
ID   Q9XYC7      PRELIMINARY;      PRT;   449 AA.
AC   Q9XYC7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HISTONE DEACETYLASE.
GN   HDAC1.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3D7;
RX   MEDLINE=99229859; PubMed=10215020;
RA   Joshi M.B., Lin D.T., Chiang P.-H., Goldman N.D., Fujioka H.,
RA   Aikiwa M., Syin C.;
RT   "Molecular cloning and nuclear localization of a histone deacetylase
RT   homologue in Plasmodium falciparum.";
RL   Mol. Biochem. Parasitol. 99:11-19(1999).
DR   EMBL; AF091326; AAD22407.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
SQ   SEQUENCE   449 AA;  51375 MW;  40F4BB4A0111AD11 CRC64;


  Query Match             50.4%;  Score 1281;  DB 5;  Length 449;
  Best Local Similarity   52.4%;  Pred. No. 9.8e-96;
  Matches  241;  Conservative   76;  Mismatches  119;  Indels   24;  Gaps    6;

Qy      18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
           :::|:||::| || |||| |||||| |||| |||| |:|:: :|: ||  :| ::: |
Db       4 RKKVAYFHDPDIGSYYYGAGHPMKPQRIRMTHSLIVSYNLYKYMEVYRPHKSDVNELTLF 63

Qy      78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGE--DCPVFDGLFDFCRASAGGSIGAAVK 135
           |  ||:|||:|:| |:  :      |:||||||  |||||||||  | :: || ||  | |
Db      64 HDYEYIDFLSSISLENYRE--FTYQLKRFNVGEATDCPVFDGLFQFQQQSCAGASIDGASK 121

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||   ||| :|| ||||||| ||||||||:|||||||||||||   ||:|||||||||||
Db     122 LNHHCADICVNWSGGLHHAKMSEASGFCYINDIVLGILELLKYHARVMYIDIDVHHGDGV 181

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           ||||| | ||||||||||||:||||| | |||   ||||::|||||||| |::|  ||:
Db     182 EEAFYVTHRVMTVSFHKFGDYFPGTGDITDVGVNHGKYYSVNVPLNDGMTDDAFVDLFKV 241

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :| | :: |:| |:::|||||||:|||| |||::|||| |:  :||||:||:|||| ||
Db     242 VIDKCVQTYRPGAIIIQCGADSLTGDRLGRFNLTIKGHARCVEHVRSYNIPLLVLGGGGY 301

Qy     316 TIRNVARCWCYETAVAVGV--EPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMER 373
           |||||:||| ||| | :    | :::  |:|::|: ||: ||: || : | |:|: : |
Db     302 TIRNVSRCWAYETGVVLNKHHEMPDQISLNDYYDYYAPDFQLHLQPSNIPNYNSPEHLSR 361

Qy     374 IRNTLLEQLSGLIHAPSVQFQHTPP--VNRVLDEPEDDMETRPKPRXWSGTATYESDSDD 431
           |:  : | |  : ||| ||| : ||   |  :|:  |  :                ||   |
Db     362 IKMKIAENLRHIEHAPGVQFSYVPPDFFNSDIDDESDKNQ----------YELKDDS 408

Qy     432 DDKPLHGYSCRGGATTD---RDSTGEDEMDDDNPEPDVNP 468
             |   :  :||    | :  :|  |  : :  : |
Db     409 GGGRAPGTRAKEHSTTHHLRRKNYDDDFFDLSDRDQSIVP 448


Search completed: April 26, 2002, 17:57:28
Job time: 356 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:57:50 ; Search time 17.07 Seconds
(without alignments)
1011.666 Million cell updates/sec

Title: US-09-645-337-4
Perfect score: 2544
Sequence: 1 MEADESGISLPSGPDGRKRR.....TGEDEMDDDNPEPDVNPPSS 471

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1561.5	61.4	501	1	HDAC_ARATH	O22446 arabidopsis
2	1517.5	59.7	513	1	HDAC_MAIZE	P56521 zea mays (m
3	1474	57.9	480	1	HDA1_CHICK	P56517 gallus gall
4	1466	57.6	482	1	HDA1_MOUSE	O09106 mus musculu
5	1465.5	57.6	488	1	HDA2_MOUSE	P70288 mus musculu
6	1462.5	57.5	482	1	HDA1_HUMAN	Q13547 homo sapien
7	1456.5	57.3	480	1	HD11_XENLA	Q91695 xenopus lae
8	1453.5	57.1	480	1	HD12_XENLA	O42227 xenopus lae
9	1453.5	57.1	520	1	HDAC_DROME	Q94517 drosophila
10	1445.5	56.8	488	1	HDA2_HUMAN	Q92769 homo sapien
11	1441.5	56.7	576	1	HDA1_STRPU	P56518 strongyloce
12	1423.5	56.0	488	1	HDA2_CHICK	P56519 gallus gall
13	1359.5	53.4	461	1	HDA1_CAEEL	O17695 caenorhabdi
14	1328.5	52.2	433	1	RPD3_YEAST	P32561 saccharomyc
15	1300.5	51.1	428	1	HDA3_CHICK	P56520 gallus gall
16	1292.5	50.8	424	1	HDA3_MOUSE	O88895 mus musculu
17	1292.5	50.8	428	1	HDA3_HUMAN	O15379 homo sapien
18	1143.5	44.9	434	1	PHD1_SCHPO	O13298 schizosacch
19	1041.5	40.9	507	1	HDA2_CAEEL	Q09440 caenorhabdi
20	1003.5	39.4	452	1	HOS2_YEAST	P53096 saccharomyc
21	436	17.1	470	1	HOS1_YEAST	Q12214 saccharomyc
22	434.5	17.1	387	1	ACUC_BACSU	P39067 bacillus su
23	394.5	15.5	385	1	ACUC_STAXY	Q56195 staphylococ
24	342	13.4	359	1	Y130_ARCFU	O30107 archaeoglob
25	309.5	12.2	706	1	HDA1_YEAST	P53973 saccharomyc
26	273.5	10.8	1113	1	HDA5_MOUSE	Q9z2v6 mus musculu
27	270.5	10.6	1080	1	HDA4_CHICK	P83038 gallus gall
28	266.5	10.5	1084	1	HDA4_HUMAN	P56524 homo sapien
29	264	10.4	687	1	HDA1_SCHPO	P56523 schizosacch
30	263.5	10.4	1122	1	HDA5_HUMAN	Q9uql6 homo sapien
31	261.5	10.3	343	1	Y535_METJA	Q57955 methanococc
32	251.5	9.9	1149	1	HDA6_MOUSE	Q9z2v5 mus musculu
33	247.5	9.7	1215	1	HDA6_HUMAN	Q9ubn7 homo sapien
34	245.5	9.7	331	1	YB94_METTH	O27262 methanobact
35	241	9.5	798	1	YLFN_CAEEL	Q20296 caenorhabdi
36	232.5	9.1	310	1	YGLA_SYNP2	P28606 synechococc
37	232	9.1	304	1	Y245_SYNY3	P72702 synechocyst
38	179	7.0	341	1	APHA_MYCRA	Q48935 mycoplana r
39	108.5	4.3	1513	1	RPOD_ORYSA	P12093 oryza sativ
40	105	4.1	639	1	NOSZ_RHIME	Q59746 rhizobium m
41	104	4.1	690	1	UBPB_HUMAN	P51784 homo sapien
42	103	4.0	347	1	MREB_ECOLI	P13519 escherichia
43	102	4.0	1375	1	NID2_HUMAN	Q14112 homo sapien
44	100.5	4.0	926	1	SC24_YEAST	P40482 saccharomyc
45	100	3.9	697	1	HOS3_YEAST	Q02959 saccharomyc

ALIGNMENTS

RESULT 1
HDAC_ARATH
ID HDAC_ARATH STANDARD; PRT; 501 AA.
AC O22446;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HISTONE DEACETYLASE (HD).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Tomihama T., Shoji K., Hanyu H., Okano T.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC FAMILY. HD SUBFAMILY 1.
CC ----------

CC ----------
DR EMBL; AF014824; AAB66486.1; -.
DR InterPro; IPR000286; His_deacetylse.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
DR PRINTS; PR01271; HISDACETLASE.
KW Hydrolase; Nuclear protein.
SQ SEQUENCE 501 AA; 56037 MW; C50AF5624958D6C2 CRC64;

Query Match 61.4%; Score 1561.5; DB 1; Length 501;
Best Local Similarity 60.7%; Pred. No. 1.7e-121;
Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;

```
Qy       4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
           |  | || ||||| ||:| |||:| :|:||||||||||||||||| |:|: || | : :::
Db       2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61

Qy      64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
             :|  |   |: |||: :|| || |::||:  |    | |:|||||||||||||: ||:
Db      62 LKPFPARERDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119
```

```
Qy     124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
           ||||:| :|||| ||||| |||||||| |||||||||||| |||||| :|||
Db     120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179

Qy     184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
           |:|||:|||||||||||| |||||||||||||:|||||||:|:|   ||||:|||||:||
Db     180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239

Qy     244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
           :||||:  ||:|:: ||||:::| |||||||||||||||||||:||||:|:::|:||:
Db     240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299

Qy     304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363
           ||||::||| |||||||||||||| ||:||| ::|:| :||:||||||||||| || ||
Db     300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359

Qy     364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXWSG 421
           | |: : :| ||| ||  || | ||||| ||  ||      :|| ::| : |  |
Db     360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDS--- 416

Qy     422 TATYESDSDDDDKP------------------LHGYSCRG-GATTDRDSTGEDEMDDDN 461
               : | ||| ||                  | |   || |   : | :|  ::   |
Db     417 ----DMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472

Qy     462 P 462
           |
Db     473 P 473
```

```
RESULT   2
HDAC_MAIZE
ID   HDAC_MAIZE     STANDARD;      PRT;   513 AA.
AC   P56521;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE (RPD3 HOMOLOG).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC   Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. W22;
RA   Rossi V., Hartings H., Motto M.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF035815; AAC50038.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   513 AA;  57546 MW;  C45387CF3A38906F CRC64;
```

```
  Query Match             59.7%;  Score 1517.5;  DB 1;  Length 513;
  Best Local Similarity   65.2%;  Pred. No. 7.7e-118;
  Matches  283;  Conservative   60;  Mismatches   72;  Indels   19;  Gaps    6;

Qy       3 ADESGISLPS-GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
           |   | |||| |||||:|||| ||||:| :|:|||||||||||||||| |||:  | |  ::
Db       6 AGSGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPMKPHRIRMTHSLLARYGLLNQM 65

Qy      62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDF 121
           :: ||: |   :: ||| :|::|| ||:||:  |    | |:|||||:||| |||: |
Db      66 QVYRPNPARERELCRFHAEEYINFLRSVTPETQQD--QIRLLKRFNVGEECPVLDGLYSF 123

Qy     122 CRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKR 181
           |:  || |:| ||| |    |||||| |||||||| |||||||||||| ||||||  :|
Db     124 CQTYAGASVGGAVKFN-HGHDIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKHHER 182

Qy     182 VLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLN 241
           |||:|||:|||||||||||||||||||||||||||:||||| |||:|  |||||:|||||:
Db     183 VLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGHSKGKYYSLNVPLD 242

Qy     242 DGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLR 301
           ||:||||::|||:|:: |||||::| ||||||||||||||||||||:||||:|:|::|
Db     243 DGIDDESYQSLFKPIMGKVMEVFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYMR 302

Qy     302 SYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP 361
           |:||||::||| |||||||||||||| ||:| ||::|:| |||:|||||||||| ||
Db     303 SFNVPLLLLGGGGYTIRNVARCWCYETGVALGQEPEDKMPVNEYYEYFGPDYTLHVAPSN 362

Qy     362 MENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXW 419
           ||| || : :: ||:    :|| | ||||| ||   |   :   || :|| : |  |
Db     363 MENKNTRQQLDDIRS----KLSKLRHAPSVHFQERVPDTEIPEQDEDQDDPDERHDP--- 415

Qy     420 SGTATYESDSDDDD 433
                 :|| : ||
Db     416 ------DSDMEVDD 423
```

```
RESULT   3
HDA1_CHICK
ID   HDA1_CHICK     STANDARD;      PRT;   480 AA.
AC   P56517;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1 OR HDAC1A.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takami Y.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sun J.M., Chen H.Y., Davie J.R.;
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: FORMS A COMPLEX WITH RBAP48 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AF039751; AAB96923.1; -.
DR   EMBL; AF043328; AAB99850.1; -.
DR   EMBL; AF044169; AAC00504.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      299     302       POLY-GLY.
FT   DOMAIN      396     399       POLY-GLU.
FT   DOMAIN      446     453       POLY-GLU.
FT   CONFLICT      3       4       LT -> VM (IN REF. 2; AAB99850).
SQ   SEQUENCE   480 AA;  54938 MW;  82C78CE285C779D9 CRC64;


  Query Match             57.9%;  Score 1474;  DB 1;  Length 480;
  Best Local Similarity   58.8%;  Pred. No. 2.8e-114;
  Matches  270;  Conservative   80;  Mismatches   95;  Indels   14;  Gaps    5;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | ||:| |:|:  :|:||||||||||||||||| |:|:::| |:|::|| ||  |:| ::
Db       6 GTKRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|: || |: |::| : |  : ::|||||||||||||||:||: |||||: :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSAGGSVASAVK 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  |||:|| |||||||||||||||||||||| |||||  :||||||||:||||||
Db     124 LNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||||||||:|::|||||  :||:|| ||||||:| || ||:||||: ::|:|
Db     184 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKP 243

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :| |||| :|| |||||||:|||||||||||||::|||| |: |::|:|:|:::||| ||
Db     244 VISKVMETFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGY 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           ||||||||| ||||||:  |  |:||||:||||||||: ||: || | | || : :|:|:
Db     304 TIRNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVL----DEPEDDMETRPKPRXWSGTATYE---SD 428
             | | |  | ||| || |  |  : |     || |:| | |   |      : :   ||
Db     364 QRLFENLRMLPHAPGVQMQPIPE-DAVQEDSGDEEEEDPEKRISIRNSDKRISCDEEFSD 422

Qy     429 SDDDD----KPLHGYSCRGGATTDRDSTGEDEMDDDNPE 463
           |:|:     | :  :       |: :   |:: |:   |
Db     423 SEDEGEGGRKNVANFKKAKRVKTEEEKEEEEKKDEKEEE 461


RESULT   4
HDA1_MOUSE
ID   HDA1_MOUSE     STANDARD;      PRT;   482 AA.
AC   O09106; P97476;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fibroblast;
RX   MEDLINE=97415582; PubMed=9271381;
RA   Bartl S., Taplick J., Lagger G., Khier H., Kuchler K., Seiser C.;
RT   "Identification of mouse histone deacetylase 1 as a growth factor-
RT   inducible gene.";
RL   Mol. Cell. Biol. 17:5033-5043(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Johnson C.A.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- FUNCTION: ACTS AS A TRANSCRIPTIONAL REPRESSOR.
CC   -!- SUBUNIT: FORMS A COMPLEX WITH RBAP48 (BY SIMILARITY). ALSO FORMS A
CC       COMPLEX WITH SIN3 AND SAP18 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGHER LEVELS IN THYMUS
CC       AND TESTIS AND LOWER LEVELS IN LIVER.
CC   -!- INDUCTION: BY INTERLEUKIN-2.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; X98207; CAA66870.1; -.
DR   EMBL; U80780; AAB68398.1; -.
DR   MGD; MGI:108086; Hdac1.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   482 AA;  55075 MW;  7F64D3C17F5E4844 CRC64;


  Query Match             57.6%;  Score 1466;  DB 1;  Length 482;
  Best Local Similarity   58.9%;  Pred. No. 1.3e-113;
  Matches  267;  Conservative   77;  Mismatches  101;  Indels    8;  Gaps    3;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | ||:| |:|:  :|:|||||||||||||||| |:|:::| |:|::|| ||  |:| ::
Db       6 GTKRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|: || |: |::| : |  : ::|||||||||||||||:||: | |||: :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVK 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  |||:|| |||||||||||||||||||||| |||||  :||||||||:||||||
Db     124 LNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||||||||:|::|||||  :||:|| ||||||:| || ||:||||: ::|:|
Db     184 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKP 243

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :: ||||::|| |||||||:|||||||||||||::|||| |: |::|:|:|:::||| ||
Db     244 VMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGY 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           ||||||||| ||||||:  |  |:||||:||||||||: ||: || | | || : :|:|:
Db     304 TIRNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPR-----XWSGTATYESDSD 430
           | | |  | ||| || |  |  : : :|  |: |  |  |             |  | ||
Db     364 QRLFENLRMLPHAPGVQMQAIPE-DAIPEESGDEDEEDPDKRISICSSDKRIACEEEFSD 422
```

```
Qy    431 DDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPE 463
          |::   |          :    ||| : |  |
Db    423 SDEEGEGGRKNSSNFKKAKRVKTEDEKEKDPEE 455


RESULT   5
HDA2_MOUSE
ID   HDA2_MOUSE     STANDARD;      PRT;   488 AA.
AC   P70288;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 2 (HD2) (YY1 TRANSCRIPTION FACTOR BINDING
DE   PROTEIN).
GN   HDAC2 OR YY1BP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lymphoma;
RX   MEDLINE=97075080; PubMed=8917507;
RA   Yang W.-M., Inouye C.J., Zeng Y., Bearss D., Seto E.;
RT   "Transcriptional repression by YY1 is mediated by interaction with a
RT   mammalian homolog of the yeast global regulator RPD3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
CC       AT LEAST WITH THE ZINC-FINGER TRANSCRIPTION FACTOR YY1.
CC   -!- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX WITH YY1.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31758; AAC52889.1; -.
DR   MGD; MGI:1097691; Hdac2.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      300    303       POLY-GLY.
SQ   SEQUENCE   488 AA;  55302 MW;  B9843D24A775157C CRC64;


  Query Match             57.6%;  Score 1465.5;  DB 1;  Length 488;
  Best Local Similarity   60.3%;  Pred. No. 1.4e-113;
  Matches  270;  Conservative   68;  Mismatches   83;  Indels   27;  Gaps    6;

Qy     16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
          | |::| |:|:  ||:|||||||||||||||| |:|:::| |:|::|| ||  | | ::
Db      7 GGKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKATAEEMT 66

Qy     76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
          ::|| ||: || |: |::| : |  : ::|||||||||||||||:||: | |||:  |||
Db     67 KYHSDEYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSTGGSVAGAVK 124

Qy    136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
          ||||  |:|:|| |||||||||||||||||||||| ||||||  :||||||||:||||||
Db    125 LNRQQTDMAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 184

Qy    196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
          |||||||||||||||||:|::||||| :||:|| |||||||:| |: ||:|||||:  :|:|
Db    185 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNFPMRDGIDDESYGQIFKP 244

Qy    256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
          :| ||||:||| ||||||||||||||||||||:||||| |:   :::|:||::||| ||
Db    245 IISKVMEMYQPSAVVLQCGADSLSGDRLGCFNLTVKGHAKCVEVAKTFNLPLLMLGGGGY 304

Qy    316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
          ||||||||| ||||||:  |  |:||||:||||||||: ||: || | | |||: ||:|:
Db    305 TIRNVARCWTYETAVALDCEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTPEYMEKIK 364

Qy    376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
          | | |  | ||| || |  |      |   :|          ||    : | :| ||
Db    365 QRLFENLRMLPHAPGVQMQAIPE-----DAVHED----------SG----DEDGEDPDKR 405

Qy    436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
          :   | |    :|:    ::|  |   |
Db    406 I---SIR---ASDKRIACDEEFSDSEDE 427


RESULT   6
HDA1_HUMAN
ID   HDA1_HUMAN     STANDARD;      PRT;   482 AA.
AC   Q13547; Q92534;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1 OR RPD3L1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=T-cell;
RX   MEDLINE=96185499; PubMed=8602529;
RA   Taunton J., Hassig C.A., Schreiber S.L.;
RT   "A mammalian histone deacetylase related to the yeast transcriptional
RT   regulator Rpd3p.";
RL   Science 272:408-411(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fetal lung;
RX   MEDLINE=96244606; PubMed=8646880;
RA   Furukawa Y., Kawakami T., Sudo K., Inazawa J., Matsumine A.,
RA   Akiyama T., Nakamura Y.;
RT   "Isolation and mapping of a human gene (RPD3L1) that is homologous to
RT   RPD3, a transcription factor in Saccharomyces cerevisiae.";
RL   Cytogenet. Cell Genet. 73:130-133(1996).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS.
CC   -!- SUBUNIT: FORMS A COMPLEX WITH RBAP48 AND ALSO WITH THE HISTONE
CC       ACETYLTRANSFERASE P/CAF AND THE ADAPTATOR PROTEIN P300. ALSO FORMS
CC       A COMPLEX WITH SIN3 AND SAP18. INTERACTS WITH TG1F.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: UBIQUITOUS, WITH HIGHER LEVELS IN HEART,
CC       PANCREAS AND TESTIS, AND LOWER LEVELS IN KIDNEY AND BRAIN.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; U50079; AAC50475.1; -.
DR   EMBL; D50405; BAA08909.1; -.
DR   MIM; 601241; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   CONFLICT    312    312       W -> R (IN REF. 2).
SQ   SEQUENCE   482 AA;  55103 MW;  4D35B7C1ED7838D6 CRC64;


  Query Match             57.5%;  Score 1462.5;  DB 1;  Length 482;
  Best Local Similarity   58.5%;  Pred. No. 2.5e-113;
  Matches  262;  Conservative   79;  Mismatches   80;  Indels   27;  Gaps    4;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | :|:| |:|:  :|:|||||||||||||||| |:|:::| |:|::|| ||  |:| ::
Db       6 GTRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|: || |: |::| : |  : ::|||||||||||||||:||: | |||: :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVK 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  |||:|| |||||||||||||||||||||| |||||  :||||||||:||||||
Db     124 LNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||||||||:|::|||| :|||:|| |||||:| || ||:|||: ::|:|
Db     184 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKP 243

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :: ||||:|| |||||||:|||||||||||||::|||| |: |::|:|:|:::||| ||
Db     244 VMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGY 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           |||||||| |||||:  |  |:||||:||||||||: ||: || | | || : :|:|:
Db     304 TIRNVARCWTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
             | |  | |||| || |  |  : : :|  |:                  | || ||
Db     364 QRLFENLRMLPHAPGVQMQAIPE-DAIPEESGDE-----------------DEDDPDKR 404

Qy     436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
           :       ::|:    |:|  |   |
Db     405 I------SICSSDKRIACEEEFSDSEEE 426


RESULT   7
HD11_XENLA
ID   HD11_XENLA     STANDARD;      PRT;   480 AA.
AC   Q91695;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE 1-1 (HD1) (MATERNALLY-EXPRESSED
DE   HISTONE DEACETYLASE) (HDM) (AB21).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oocyte;
RX   MEDLINE=98036059; PubMed=9370292;
RA   Ladomery M.R., Lyons S., Sommerville J.;
RT   "Xenopus HDm, a maternally expressed histone deacetylase, belongs to
RT   an ancient family of acetyl-metabolizing enzymes.";
RL   Gene 198:275-280(1997).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- TISSUE SPECIFICITY: OOCYTE.
CC   -!- DEVELOPMENTAL STAGE: ACCUMULATES IN PREVITELLOGENIC OOCYTES AND IS
CC       MAINTAINED AT CONSTANT LEVEL THROUGHOUT OOGENESIS AND INTO EARLY
CC       EMBRYOGENESIS. DECLINES THROUGH GASTRULA TO NEURULA. NOT
CC       DETECTABLE BETWEEN NEURULA AND TAILBUD, NOR IN ADULT TISSUES OTHER
CC       THAN OVARY.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X78454; CAA55211.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      299    302       POLY-GLY.
SQ   SEQUENCE   480 AA;  54747 MW;  7B831822235DADB5 CRC64;


  Query Match             57.3%;  Score 1456.5;  DB 1;  Length 480;
  Best Local Similarity   59.4%;  Pred. No. 7.7e-113;
  Matches  266;  Conservative   75;  Mismatches   80;  Indels   27;  Gaps    6;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | |::| |:|:  :|:|||||||||||||||| |:|:::| |:|::|| ||  | | |:
Db       6 GTKKKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIFRPHKASAEDMT 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|: || |: |::| : |  : ::|||||||||||||||:||: |||||: :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSAGGSVASAVK 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  ||::|| |||||||||||||||||||| ||||||  :||:|||||:||||||
Db     124 LNKQQTDISVNWSGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVVYIDIDIHHGDGV 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||||||||:|::|||| :||:|| ||||||:|  | ||:|||||: ::|:|
Db     184 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYALRDGIDDESYEAIFKP 243

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :: ||||::|| |||||||||||||||||||::|||| |: |::::|:||::|||: ::|:|
Db     244 VMSKVMEMFQPSAVVLQCGADSLSGDRLGCFNLTIKGHAKCVEFIKTFNLPLLMLGGGGY 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           |||||||| |||||:  |  |:||||:||||||||: ||: || | | || : :|:|:
Db     304 TIRNVARCWTYETAVALDSEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
             | |  | ||| || |           |   ||           ||    | | || ||
Db     364 QRLFENLRMLPHAPGVQMQAVAE-----DSIHDD----------SG----EEDEDDPDKR 404

Qy     436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
           :   | |   ::|:    ::|  |   |
Db     405 I---SIR---SSDKRIACDEEFSDSEDE 426


RESULT   8
HD12_XENLA
ID   HD12_XENLA     STANDARD;      PRT;   480 AA.
AC   O42227;
```

```
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE 1-2 (HD1) (RPD3 HOMOLOG).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Patterton D., Wolffe A.P.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF020658; AAC60346.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      299     302       POLY-GLY.
SQ   SEQUENCE   480 AA;  54892 MW;  CA92DE34D36E39E8 CRC64;


  Query Match             57.1%;  Score 1453.5;  DB 1;  Length 480;
  Best Local Similarity   58.7%;  Pred. No. 1.4e-112;
  Matches  263;  Conservative   79;  Mismatches   79;  Indels   27;  Gaps    6;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | |::| |:|:  :|:|||||||||||||||| |:|:::| |:|::|| ||  | | ::
Db       6 GTKKKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKASAEEMT 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|: || |: |::| : |  : ::||||||||||||||||:||: | |||: :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVK 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  ||::|| |||||||||||||||||||||| ||||||  :||:|||||:|||||
Db     124 LNKQQTDISVNWSGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVVYIDIDIHHGDGV 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||:|||||:|::||||| :||:|| ||||||:| || ||:|||||: ::|:|
Db     184 EEAFYTTDRVMSVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKP 243

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :: ||||::|| |||||||||||||||||||||::|||| |: |::::|:|:::||| ||
Db     244 VMTKVMEMFQPSAVVLQCGADSLSGDRLGCFNLTIKGHAKCVEFIKTFNLPMLMLGGGGY 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           |||||||| |||||:  |  |:||||:||||||||: ||: || | | || : :|:|:
Db     304 TIRNVARCWTYETAVALDSEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
           | | |  | ||| || |  |      |  ||          ||    | | :| ||
Db     364 QRLFENLRMLPHAPGVQMQAIPE-----DSVHDD----------SG----EEDEEDPDKR 404

Qy     436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
```

```
           :  | |  ::|:    ::|  |  |
Db     405 I---SIR---SSDKRIACDEEFSDSEDE 426


RESULT   9
HDAC_DROME
ID   HDAC_DROME     STANDARD;      PRT;   520 AA.
AC   Q94517; O17429;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE (HD) (DRPD3).
GN   RPD3 OR HDAC1.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   MEDLINE=97113416; PubMed=8955276;
RA   de Rubertis F., Kadosh D., Henchoz S., Pauli D., Reuter G., Struhl K.,
RA   Spierer P.;
RT   "The histone deacetylase RPD3 counteracts genomic silencing in
RT   Drosophila and yeast.";
RL   Nature 384:589-591(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Johnson C.A., White D., O'Neill L.P., Turner B.M.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY). IN DROSOPHILA, IT IS INVOLVED IN POSITION-EFFECT
CC       VARIEGATION (PEV) AND IS DIRECTLY INVOLVED IN COUNTERACTING
CC       GENOME SILENCING.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y09258; CAA70455.1; -.
DR   EMBL; AF026949; AAC23917.1; -.
DR   FlyBase; FBgn0015805; Rpd3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   CONFLICT     50     51       DI -> EIY (IN REF. 2).
FT   CONFLICT     66     66       C -> S (IN REF. 2).
FT   CONFLICT     96     96       D -> N (IN REF. 2).
FT   CONFLICT    105    105       E -> D (IN REF. 2).
FT   CONFLICT    295    295       V -> VV (IN REF. 2).
FT   CONFLICT    370    370       L -> V (IN REF. 2).
FT   CONFLICT    506    506       S -> T (IN REF. 2).
SQ   SEQUENCE   520 AA;  58183 MW;  D02EA3DBD3C64688 CRC64;


  Query Match             57.1%;  Score 1453.5;  DB 1;  Length 520;
  Best Local Similarity   57.9%;  Pred. No. 1.5e-112;
  Matches  270;  Conservative   77;  Mismatches   90;  Indels   29;  Gaps    6;

Qy      18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
```

```
          |:|| |:|:  ||:|||||||||||||||| |:|:::| |:|:::| ||  | | :: :|
Db      6 KKRVCYYYDSDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMDI-RPHKATADEMTKF 64

Qy     78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
          |  ||| || |: |::| : :  : ::|||||||||||||||::||: |||||: ||||||
Db     65 HCDEYVRFLRSIRPDNMSEYN--KQMQRFNVGEDCPVFDGLYEFCQLSAGGSVAAAVKLN 122

Qy    138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEE 197
          :| ::| |||||||||||||||||||||||||||||||||  :|||||||||||||||||
Db    123 KQASEICINWGGGLHHAKKSEASGFCYVNDIVLGILELLKYHQRVLYIDIDVHHGDGVEE 182

Qy    198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
          |||||||||||||||:|::||||| :||:|| |||||:|:|| |||||::: |:| |:|
Db    183 AFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNIPLRDGMDDDAYESIFVPII 242

Qy    258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
          |||| :|| ||||||||||:|||||||||:|||| |: |:: ||:| :::|| ||||
Db    243 SKVMETFQPAAVVLQCGADSLTGDRLGCFNLTVKGHGKCVEFVKKYNLPFLMVGGGGYTI 302

Qy    318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNT 377
          | |:||| |||:|| ||  |:||||:||||||||: ||: || | | || : :|:|:|
Db    303 RKVSRCWTYETSVALAVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTSEYLEKIKNR 362

Qy    378 LLEQLSGLIHAPSVQFQHTP--PVNRVLDE----------PEDDMETRPKPRXWSGTATY 425
          | | |  | |||| || |  |   :|   |:          |: | : |  |
Db    363 LFENLRMLPHAPGVQIQAIPEDAINDESDDEDKVDKDDRLPQSDKDKRIVPE------NE 416

Qy    426 ESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
           |||:|: :        ||   :|   |: :    : : :  |  ||
Db    417 YSDSEDEGE--------GGRRDNRSYKGQRKRPRLDKDTNSNKASS 454


RESULT  10
HDA2_HUMAN
ID   HDA2_HUMAN     STANDARD;      PRT;   488 AA.
AC   Q92769;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 2 (HD2).
GN   HDAC2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Breast;
RX   MEDLINE=97075080; PubMed=8917507;
RA   Yang W.-M., Inouye C.J., Zeng Y., Bearss D., Seto E.;
RT   "Transcriptional repression by YY1 is mediated by interaction with a
RT   mammalian homolog of the yeast global regulator RPD3.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:12845-12850(1996).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- FUNCTION: FORMS TRANSCRIPTIONAL REPRESSOR COMPLEXES BY ASSOCIATING
CC       WITH MAD, SIN3, YY1 AND N-COR.
CC   -!- SUBUNIT: FORMS A HETEROLOGOUS COMPLEX AT LEAST WITH YY1.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; LOWER LEVEL IN BRAIN AND
CC       LUNG.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U31814; AAC50814.1; -.
DR   MIM; 605164; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      300    303       POLY-GLY.
SQ   SEQUENCE   488 AA;  55325 MW;  3AE554ADC7734B70 CRC64;


  Query Match             56.8%;  Score 1445.5;  DB 1;  Length 488;
  Best Local Similarity   59.6%;  Pred. No. 6.4e-112;
  Matches  267;  Conservative   68;  Mismatches   86;  Indels   27;  Gaps    6;

Qy     16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
          | |::| |:|:  ||:|||||||||||||||| |:|:::| |:|:::|| ||  | | ::
Db      7 GGKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKATAEEMT 66

Qy     76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
          ::|| ||: || |: |::| : |  : :  |||||||| |||||:||: | |||:  |||
Db     67 KYHSDEYIKFLRSIRPDNMSEYS--KQMHIFNVGEDCPAFDGLFEFCQLSTGGSVAGAVK 124

Qy    136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
          ||||  |:|:|| |||||||| |||||||||||| ||||||  :|||||||:||||||
Db    125 LNRQQTDMAVNWAGGLHHAKKYEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 184

Qy    196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
          ||||||||||||||||:|::||||| :||:|| ||||||:| |: ||:|||||:  :|:|
Db    185 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNFPMCDGIDDESYGQIFKP 244

Qy    256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
          :| ||||:||| ||||||||||||||||||||:||||| |:  ::::|:||::||| ||
Db    245 IISKVMEMYQPSAVVLQCGADSLSGDRLGCFNLTVKGHAKCVEVVKTFNLPLLMLGGGGY 304

Qy    316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
          ||||||||| ||||||:  |  |:||||:|||||||: ||: || | | |||: ||:|:
Db    305 TIRNVARCWTYETAVALDCEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTPEYMEKIK 364

Qy    376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
             | | |  | ||| || |  |      |   :|           ||    : | :| ||
Db    365 QRLFENLRMLPHAPGVQMQAIPE-----DAVHED----------SG----DEDGEDPDKR 405

Qy    436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
          :   | |    :|:    ::|  |   |
Db    406 I---SIR---ASDKRIACDEEFSDSEDE 427


RESULT  11
HDA1_STRPU
ID   HDA1_STRPU     STANDARD;      PRT;   576 AA.
AC   P56518;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 1 (HD1).
GN   HDAC1.
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nemer M.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
```

```
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF032919; AAB87685.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   576 AA;  64078 MW;  B3D11A844A2088E9 CRC64;


  Query Match             56.7%;  Score 1441.5;  DB 1;  Length 576;
  Best Local Similarity   57.2%;  Pred. No. 1.7e-111;
  Matches  267;  Conservative   80;  Mismatches  103;  Indels   17;  Gaps    6;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | |:|| |:|:  :|:|||||||||||||||| |:||::| |:|::|| ||  |   ::
Db       5 GTKKRVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLILNYGLYRKMEIYRPHKAVMEEMT 64

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|| || :: |::| :    : ::|||||||||||||::||: |:|||:  |||
Db      65 KYHSDDYVKFLRTIRPDNMSE--YTKQMQRFNVGEDCPVFDGLYEFCQLSSGGSVAGAVK 122

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  |||||| |||||||||||||||||||||| |||||  :||||||||:||||||
Db     123 LNKQQTDIAINWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 182

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||||||||:|::|||||| :||:|| |||||:| || ||:||||:  :|:|
Db     183 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNFPLRDGIDDESYDKIFKP 242

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :: ||||:||| |: ||||||||||||||||::|||| |: |:: ||:||:::|| ||
Db     243 IMCKVMEMYQPSAICLQCGADSLSGDRLGCFNLTLKGHAKCVEFMKQYNLPLLLMGGGGY 302

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           |||||||| |||: |:|||  |:||||:||||||||: ||: || | | || : :::|:
Db     303 TIRNVARCWTYETSTALGVEIANELPYNDYFEYFGPDFKLHISPSNMTNQNTGEYLDKIK 362

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYE--------S 427
           | | :  : ||| || |  |  : : |: : : |     :  |  |  :         |
Db     363 TRLYENMRMIPHAPGVQMQPIPE-DAIPDDSDAEDEAENPDKRISIMAQDKRIQRDDEFS 421

Qy     428 DSDDD-DKPLHGYSCRGGATTDRDSTGEDEMDDDNP--EPDVNPPSS 471
           ||:|: :  | |    ||    :    ||:|  | |    ||
Db     422 DSEDEGETRLPG---EGGRRDHRSHKAKRSKIDDSPGKEADKEAKSS 465


RESULT  12
HDA2_CHICK
ID   HDA2_CHICK     STANDARD;      PRT;   488 AA.
AC   P56519;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 2 (HD2).
GN   HDAC2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takami Y.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039752; AAB96924.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
FT   DOMAIN      300    303       POLY-GLY.
SQ   SEQUENCE   488 AA;  55153 MW;  4F79B9C0D4A2D065 CRC64;


  Query Match             56.0%;  Score 1423.5;  DB 1;  Length 488;
  Best Local Similarity   59.4%;  Pred. No. 4.2e-110;
  Matches  266;  Conservative   67;  Mismatches   88;  Indels   27;  Gaps    6;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | |::| |:|:  ||:|||||||||||||||| |:|:::| |:|::|| ||  | | ::
Db       7 GGKKKVCYYYDGDIGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKATAEEMT 66

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| ||: || |: |::| : |  : ::|||||||||||||||:||: | |||:  |||
Db      67 KYHSDEYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSTGGSVAGAVK 124

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           |||| |:|:|| ||||||||||||||||||||| |||||  :|||||||:||||||
Db     125 LNRQQTDMAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 184

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           ||||||||||||||     : ||||| :||:|| ||||||:| |: ||:|||||:  :|:|
Db     185 EEAFYTTDRVMTVSEVSMVNNFPGTGDLRDIGAGKGKYYAVNFPMRDGIDDESYGQIFKP 244

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :| ||||:||| ||||||||||||||||||||:||||| |:  ::::|:||::||| ||
Db     245 IISKVMEMYQPSAVVLQCGADSLSGDRLGCFNLTVKGHAKCVEVVKTFNLPLLMLGGGGY 304

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           |||||||| ||||||:  |  |:||||:||||||||: ||: || | | ||||: ||:|:
Db     305 TIRNVARCWTYETAVALDCEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTPEYMEKIK 364

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
           | | |  | ||| || |  |      |   :|          ||    : | :| ||
Db     365 QRLFENLRMLPHAPGVQMQAIPE-----DAVHED----------SG----DEDGEDPDKR 405

Qy     436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
           :   | |    :|:    ::|  |   |
Db     406 I---SIR---ASDKRIACDEEFSDSEDE 427


RESULT  13
HDA1_CAEEL
ID   HDA1_CAEEL     STANDARD;      PRT;   461 AA.
AC   O17695;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
```

```
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PROBABLE HISTONE DEACETYLASE 1.
GN   HDA-1 OR C53A5.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Mortimore B.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z81486; CAB03984.1; -.
DR   WormPep; C53A5.3; CE08952.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   461 AA;  52137 MW;  9A104E88C5A1C07A CRC64;


  Query Match             53.4%;  Score 1359.5;  DB 1;  Length 461;
  Best Local Similarity   57.3%;  Pred. No. 7.6e-105;
  Matches  262;  Conservative   65;  Mismatches  109;  Indels   21;  Gaps    5;

Qy      12 SGP--DGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLA 69
           :||  :  |||||:|:|:  ||:|||||||| |||||||| | |:::| |:| ||| ||  |
Db       4 NGPLMEHGKRRVAYYYDSNIGNYYYGQGHVMKPHRIRMTHHLVLNYGLYRNLEIFRPFPA 63

Qy      70 DASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGS 129
             |: |||| ||: || | :|:::   |  : : :|||||||||:||||::||: |:|||
Db      64 SFEDMTRFHSDEYMTFLKSANPDNL--KSFNKQMLKFNVGEDCPLFDGLYEFCQLSSGGS 121

Qy     130 IGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDV 189
           : || |||:|  |||||| ||||||||||||||||| ||||||||||||  |||||:||||
Db     122 LAAATKLNKQKVDIAINWMGGLHHAKKSEASGFCYTNDIVLGILELLKYHKRVLYVDIDV 181

Qy     190 HHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESF 249
           |||||||||||||||||||||||:|||||||| ::|:|| ||| |::|||| ||: | |:
Db     182 HHGDGVEEAFYTTDRVMTVSFHKYGDFFPGTGDLKDIGAGKGKLYSVNVPLRDGITDVSY 241

Qy     250 RSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMV 309
           :|:|:|:: |||| : | ||||||||||:||||| |||::||| :| || ||||||||:
Db     242 QSIFKPIMTKVMERFDPCAVVLQCGADSLNGDRLGPFNLTLKGHGECARFFRSYNVPLMM 301

Qy     310 LGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPK 369
           :|| ||| ||||||| |||::||  |  |:||||:||||||:| ||:: |   | |:
Db     302 VGGGGYTPRNVARCWTYETSIAVDKEVPNELPYNDYFEYFGPNYRLHIESSNAANENSSD 361

Qy     370 DMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDS 429
           : :::  ::  |  |  |||| :  |     |    ||              :  :|
Db     362 MLAKLQTDVIANLEQLTFVPSVQMRPIP--EDALSALNDD-------------SLIADQ 405

Qy     430 DDDDKPLHGYSCRGGATTDRD-STGEDEMDDDNPEPD 465
           : || |      |    | |   || | ||   | |
Db     406 ANPDKRLPPQITDGMIQDDGDFYDGEREGDDRRNESD 442


RESULT  14
RPD3_YEAST
ID   RPD3_YEAST     STANDARD;      PRT;   433 AA.
AC   P32561;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE RPD3 (TRANSCRIPTIONAL REGULATORY PROTEIN RPD3).
GN   RPD3 OR SDI2 OR YNL330C OR N0305.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92049362; PubMed=1944291;
RA   Vidal M., Gaber R.F.;
RT   "RPD3 encodes a second factor required to achieve maximum positive
RT   and negative transcriptional states in Saccharomyces cerevisiae.";
RL   Mol. Cell. Biol. 11:6317-6327(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RA   van Dyck L., Pascual-Ahuir A., Goffeau A.;
RL   Submitted (DEC-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 1-201 FROM N.A.
RC   STRAIN=S288C / FY1676;
RX   MEDLINE=95373280; PubMed=7645347;
RA   Maftahi M., Nicaud J., Levesque H., Gaillardin C.;
RT   "Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV
RT   identifies the RPD3, PAS8 and KRE1 loci, five new open reading
RT   frames.";
RL   Yeast 11:567-572(1995).
RN   [4]
RP   CHARACTERIZATION.
RX   MEDLINE=97121415; PubMed=8962081;
RA   Rundlett S.E., Carmen A.A., Kobayashi R., Bavykin S., Turner B.M.,
RA   Grunstein M.;
RT   "HDA1 and RPD3 are members of distinct yeast histone deacetylase
RT   complexes that regulate silencing and transcription.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:14503-14508(1996).
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION AND CELL CYCLE PROGRESSION. REQUIRED FOR BOTH FULL
CC       TRANSCRIPTION REPRESSION AND ACTIVATION OF MANY GENES INCLUDING
CC       CELL TYPE-SPECIFIC GENES (STE6, TY2 AND HO), CELL DIFFERENTIATION-
CC       SPECIFIC GENES (SPO13), GENES THAT RESPOND TO EXTERNAL SIGNALS
CC       (PHO5) AND TRK2.
CC   -!- SUBUNIT: IT PROBABLY FORMS A COMPLEX WITH THE PAIRED AMPHIPATHIC
CC       HELIX PROTEIN SIN3/RPD1. INTERACTS WITH CYCLOPHILINS CPR6 AND
CC       CPR7.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S66438; AAB20328.1; -.
DR   EMBL; X83226; CAA58228.1; -.
DR   EMBL; Z46259; CAA86368.1; -.
DR   EMBL; Z71605; CAA96262.1; -.
DR   EMBL; Z71606; CAA96263.1; -.
```

```
DR   PIR; S22284; S22284.
DR   SGD; S0005274; RPD3.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   433 AA;  48904 MW;  34FFD72A7E7425DB CRC64;


  Query Match             52.2%;  Score 1328.5;  DB 1;  Length 433;
  Best Local Similarity   57.9%;  Pred. No. 2.5e-102;
  Matches  238;  Conservative   74;  Mismatches   94;  Indels    5;  Gaps    2;

Qy      18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
           ||||:|||:  :|:| || ||||||||||||||||::| |::::|| |   |   :: :|
Db      18 KRRVAYFYDADVGNYAYGAGHPMKPHRIRMAHSLIMNYGLYKKMEIYRAKPATKQEMCQF 77

Qy      78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
           |: ||:|||: |:|:::      |   :||||:||||||||:::|  | |||:  | :||
Db      78 HTDEYIDFLSRVTPDNL--EMFKRESVKFNVGDDCPVFDGLYEYCSISGGGSMEGAARLN 135

Qy     138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEE 197
           |   |:|:|: ||||||||||||||||:||||||:|||:   |||||||||||||||||
Db     136 RGKCDVAVNYAGGLHHAKKSEASGFCYLNDIVLGIIELLRYHPRVLYIDIDVHHGDGVEE 195

Qy     198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
           |||||||||| ||||:|:|||||| :||:|   || ||:|||| ||:|| ::||:| |:|
Db     196 AFYTTDRVMTCSFHKYGEFFPGTGELRDIGVGAGKNYAVNVPLRDGIDDATYRSVFEPVI 255

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
           :|:|| ||| ||||||| ||||||||||||||::|||:|: :::|: :|:||:|| |||:
Db     256 KKIMEWYQPSAVVLQCGGDSLSGDRLGCFNLSMEGHANCVNYVKSFGIPMMVVGGGGYTM 315

Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNT 377
           ||||| ||:|| :   |  |  |||||||:||:|||| | | || | |:|||: ::::
Db     316 RNVARTWCFETGLLNNVVLDKDLPYNEYYEYYGPDYKLSVRPSNMFNVNTPEYLDKVMTN 375

Qy     378 LLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESD 428
           :   |    :|||||  |||     | : |:|     : :   | : |  |
Db     376 IFANLENTKYAPSVQLNHTPRDAEDLGDVEED---SAEAKDTKGGSQYARD 423


RESULT  15
HDA3_CHICK
ID   HDA3_CHICK     STANDARD;      PRT;   428 AA.
AC   P56520;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HISTONE DEACETYLASE 3 (HD3).
GN   HDAC3.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Takami Y.;
RL   Submitted (DEC-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: RESPONSIBLE FOR THE DEACETYLATION OF LYSINE RESIDUES ON
CC       THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).
CC       HISTONE DEACETYLATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL
CC       REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: BELONGS TO THE HISTONE DEACETYLASE / ACUC / APHA
CC       FAMILY. HD SUBFAMILY 1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF039753; AAB96925.1; -.
DR   InterPro; IPR000286; His_deacetylse.
DR   Pfam; PF00850; Hist_deacetyl; 1.
DR   PRINTS; PR01270; HDASUPER.
DR   PRINTS; PR01271; HISDACETLASE.
KW   Hydrolase; Nuclear protein.
SQ   SEQUENCE   428 AA;  48901 MW;  42E32733AD2BBF07 CRC64;


  Query Match             51.1%;  Score 1300.5;  DB 1;  Length 428;
  Best Local Similarity   55.6%;  Pred. No. 5.1e-100;
  Matches  237;  Conservative   78;  Mismatches   92;  Indels   19;  Gaps    7;

Qy      19 RRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFH 78
           : |:|||:| :|:::|| |||||||: : ||||::|| |:::: : :|  |   |: |||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      79 SPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNR 138
           | :|:|||  ||| :|      ::|  ||||:||||| |||:||    | |:  | :||
Db      63 SEDYIDFLQRVSPNNM--QGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNN 120

Qy     139 QDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEA 198
           :  |||||| |||||||| ||||||||||||:|||||||   ||||||||:|||||:||
Db     121 KICDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEA 180

Qy     199 FYTTDRVMTVSFHKFGD-FFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
           || |||||||||||:|: |||||| : :|||| |:|||||||| ||:||:|:: ||:|:|
Db     181 FYLTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYALNVPLRDGIDDQSYKHLFQPVI 240

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
           :|:: |||  :|||||||||  ||||||||||::|| :|: :::|:|:||:|||| |||:
Db     241 NQVVDYYQPTCIVLQCGADSLGRDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTV 300

Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP-MENLNTPKDMERIRN 376
           ||||||| |||:: |       :|||:|||||| ||:||| | |  :|| |: : :::||
Db     301 RNVARCWTYETSLLVDEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQ 360

Qy     377 TLLEQLSGLIHAPSVQFQHTP----PVNRVLDEPE------DDMETRPKPRXWSGTATYE 426
           |: | |  | ||||||    |      :|  ||:       ::  :||:    :    |:
Db     361 TIFENLKMLNHAPSVQIHDVPSDLLSYDRT-DEPDPEERGSEENYSRPE----AANEFYD 415

Qy     427 SDSDDD 432
             | |:|
Db     416 GDHDND 421


Search completed: April 26, 2002, 17:57:52
Job time: 335 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:52:45 ; Search time 25.88 Seconds
(without alignments)
1386.330 Million cell updates/sec

Title: US-09-645-337-4
Perfect score: 2544
Sequence: 1 MEADESGISLPSGPDGRKRR.....TGEDEMDDDNPEPDVNPPSS 471

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1560.5	61.3	501	2	T05640	histone deacetylas
2	1517.5	59.7	513	2	T01413	probable histone d
3	1456.5	57.3	480	1	S60381	RPD3 protein homol
4	1359.5	53.4	461	2	T20163	hypothetical prote
5	1328.5	52.2	433	1	S22284	transcription regu
6	1305.5	51.3	465	2	T23963	hypothetical prote
7	1292.5	50.8	424	2	JC7102	histone deacetylas
8	1285.5	50.5	428	2	JC5834	histone deacetylas
9	1275	50.1	405	2	T40300	histone deacetylas
10	1207	47.4	419	2	T47443	hypothetical prote
11	1143.5	44.9	434	2	T11643	histone deacetylas
12	1041.5	40.9	507	2	T19067	hypothetical prote
13	1003.5	39.4	452	2	S64211	probable transcrip
14	457.5	18.0	375	2	D70388	acetoin utilizatio
15	436	17.1	470	2	S54089	hypothetical prote
16	434.5	17.1	387	1	S39643	acetate / acetoin
17	424.5	16.7	367	2	T36278	hypothetical prote
18	410.5	16.1	389	2	E84054	acetoin dehydrogen
19	342	13.4	359	1	B69266	acetylpolyamine am
20	319.5	12.6	380	2	D83174	probable acetylpol
21	309.5	12.2	706	2	S62933	hypothetical prote
22	305	12.0	310	1	A70481	acetoin utilizatio
23	299.5	11.8	158	2	T48929	hypothetical prote
24	269.5	10.6	425	2	G86217	protein T27G7.14 [
25	269.5	10.6	796	2	T32425	hypothetical prote
26	268	10.5	338	2	H84173	acetoin utilizatio
27	265	10.4	878	2	T17245	hypothetical prote
28	264	10.4	687	2	T43797	probable histone d
29	261.5	10.3	343	1	G64366	acetylpolyamine am
30	255.5	10.0	306	2	D82126	histone deacetylas
31	251.5	9.9	335	1	H71071	hypothetical prote
32	251.5	9.9	1095	2	T13964	probable histone d
33	245.5	9.7	331	1	C69026	acetylpolyamine am
34	243.5	9.6	385	2	A81926	hypothetical prote
35	239.5	9.4	142	2	T47441	hypothetical prote
36	239.5	9.4	1063	2	T46284	hypothetical prote
37	235.5	9.3	369	2	F81178	histone deacetylas
38	232.5	9.1	310	2	B47050	glnA 3'-region hyp
39	232	9.1	304	1	S74557	acetylpolyamine am
40	231.5	9.1	334	2	B75095	probable histone d
41	229	9.0	517	2	T27101	hypothetical prote
42	215.5	8.5	301	2	H75470	histone deacetylas
43	215.5	8.5	359	2	T05998	hypothetical prote
44	203	8.0	346	2	H83469	acetylpolyamine am
45	203	8.0	782	2	T22134	hypothetical prote

ALIGNMENTS

RESULT 1
T05640
histone deacetylase (EC 3.5.1.-) F20D10.250 - Arabidopsis thaliana
N;Alternate names: protein F20D10.250
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 12-Nov-1999
C;Accession: T05640
R;Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15420
A;Accession: T05640
A;Molecule type: DNA
A;Residues: 1-501 <BEV>
A;Cross-references: EMBL:AL035538; GSPDB:GN00062; ATSP:F20D10.250
A;Experimental source: cultivar Columbia; BAC clone F20D10
C;Genetics:
A;Gene: ATSP:F20D10.250
A;Map position: 4
A;Introns: 176/3; 323/3; 400/3; 424/2; 443/3; 467/3
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase
F;21-320/Domain: RPD3/acuC homology <RAH>

Query Match 61.3%; Score 1560.5; DB 2; Length 501;
Best Local Similarity 60.7%; Pred. No. 2.8e-116;
Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;

```
Qy       4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
           |  | || ||||| ||:| |||:| :|:|||||||||||||||| |:|: || | : :::
Db       2 DTGGNSLASGPDGVKRKVCYFYDPEVGNYYYGQGHPMKPHRIRMTHALLAHYGLLQHMQV 61

Qy      64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
             :|  |   |: |||: :|| || |::||:  |    | |:||||||||||||||: ||:
Db      62 LKPFPARDRDLCRFHADDYVSFLRSITPETQQD--QIRQLKRFNVGEDCPVFDGLYSFCQ 119

Qy     124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
             ||||:| :|||    |||||| |||||||| |||||||||||||| ||||||  :|||
Db     120 TYAGGSVGGSVKLNHGLCDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKQHERVL 179

Qy     184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
           |:|||:|||||||||||| |||||||||||||:|||||||:|:|   ||||:|||||:||
Db     180 YVDIDIHHGDGVEEAFYATDRVMTVSFHKFGDYFPGTGHIQDIGYGSGKYYSLNVPLDDG 239

Qy     244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
           :||||:  ||:|:: ||||:::| |||||||||||||||||||||:||||:|:::|:||:
Db     240 IDDESYHLLFKPIMGKVMEIFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMRSF 299

Qy     304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363
           |||::||| |||||||||||||| ||:||| ::|:| :||:|||||||||| || ||
Db     300 NVPLLLLGGGGYTIRNVARCWCYETGVALGVEVEDKMPEHEYYEYFGPDYTLHVAPSNME 359
```

```
Qy    364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXWSG 421
          | |: : :| ||| ||  || | ||||| ||  ||      :|| ::| : |  |
Db    360 NKNSRQMLEEIRNDLLHNLSKLQHAPSVPFQERPPDTETPEVDEDQEDGDKRWDPDS--- 416

Qy    422 TATYESDSDDDDKP------------------LHGYSCRG-GATTDRDSTGEDEMDDDN 461
              : | ||| ||                  | |   || |   : | :|  ::   |
Db    417 ----DMDVDDDRKPIPSRVKREAVEPDTKDKDGLKGIMERGKGCEVEVDESGSTKVTGVN 472

Qy    462 P 462
          |
Db    473 P 473


RESULT   2
T01413
probable histone deacetylase (EC 3.5.1.-) rpd3 - maize
C;Species: Zea mays (maize)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 12-Nov-1999
C;Accession: T01413
R;Rossi, V.; Hartings, H.; Motto, M.
Mol. Gen. Genet. 258, 288-296, 1998
A;Title: Identification and characterisation of an RPD3 homologue from maize (Zea mays L
A;Reference number: Z14321; MUID:98307342
A;Accession: T01413
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-513 <ROS>
A;Cross-references: EMBL:AF035815; NID:g2665839; PIDN:AAC50038.1; PID:g2665840
A;Experimental source: strain W22
C;Genetics:
A;Gene: rpd3
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase
F;27-325/Domain: RPD3/acuC homology <RAH>


  Query Match             59.7%;  Score 1517.5;  DB 2;  Length 513;
  Best Local Similarity   65.2%;  Pred. No. 7.6e-113;
  Matches  283;  Conservative   60;  Mismatches   72;  Indels   19;  Gaps    6;

Qy      3 ADESGISLPS-GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
          |   | |||| ||||:|||| |||:| :|:|||||||||||||||| |||:  | |  ::
Db      6 AGSGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPMKPHRIRMTHSLLARYGLLNQM 65

Qy     62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDF 121
          :: ||: |   :: |||: ||::|| ||:||:  |    | |:|||||:||| |||: |
Db     66 QVYRPNPARERELCRFHAEEYINFLRSVTPETQQD--QIRLLKRFNVGEECPVLDGLYSF 123

Qy    122 CRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKR 181
          |:  || |:| ||| |    |||||| |||||||| |||||||||||| ||||||  :|
Db    124 CQTYAGASVGGAVKFN-HGHDIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKHHER 182

Qy    182 VLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLN 241
          |||:|||:|||||||||||||||||||||||||:||||| |||:|  |||||:|||||:
Db    183 VLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGHSKGKYYSLNVPLD 242

Qy    242 DGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLR 301
          ||:|||||:|||:|:: |||||:|| ||||||||||||||||||||:||||:|:|::|
Db    243 DGIDDESYQSLFKPIMGKVMEVFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYMR 302

Qy    302 SYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP 361
          |:||||::||| |||||||||||||| ||:| ||::|:| |||:|||||||||| ||
Db    303 SFNVPLLLLGGGGYTIRNVARCWCYETGVALGQEPEDKMPVNEYYEYFGPDYTLHVAPSN 362

Qy    362 MENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXW 419
          ||| || : :: ||:    :|| | ||||| ||   |   :   || :|| : |  |
Db    363 MENKNTRQQLDDIRS----KLSKLRHAPSVHFQERVPDTEIPEQDEDQDDPDERHDP--- 415

Qy    420 SGTATYESDSDDDD 433
                :|| : ||
Db    416 ------DSDMEVDD 423


RESULT   3
S60381
RPD3 protein homolog - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: S60381; S43160
R;Ladomery, M.R.; Lyons, S.; Sommerville, J.
submitted to the EMBL Data Library, December 1994
A;Reference number: S60381
A;Accession: S60381
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-480 <LAD>
A;Cross-references: EMBL:X78454; NID:g773397; PIDN:CAA55211.1; PID:g602098
C;Superfamily: RPD3 protein; RPD3/acuC homology
F;13-312/Domain: RPD3/acuC homology <RAH1>


  Query Match             57.3%;  Score 1456.5;  DB 1;  Length 480;
  Best Local Similarity   59.4%;  Pred. No. 4.9e-108;
  Matches  266;  Conservative   75;  Mismatches   80;  Indels   27;  Gaps    6;

Qy     16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
          | |::| |:|:  :|:|||||||||||||||| |:|:::| |:|::|| ||  | | |:
Db      6 GTKKKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIFRPHKASAEDMT 65

Qy     76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
          ::|| :|: || |: |::| : |  : ::|||||||||||||||:|: |||||: :|||
Db     66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSAGGSVASAVK 123

Qy    136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
          ||:|  ||::|| |||||||||||||||||||||||| |||||  :||:|||||:||||||
Db    124 LNKQQTDISVNWSGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVVYIDIDIHHGDGV 183

Qy    196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
          |||||||||||||||||:|::|||||  :||:|| ||||||:|  | ||:||||: ::|:|
Db    184 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYALRDGIDDESYEAIFKP 243

Qy    256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
          :: ||||::|| |||||||||||||||||||::|||| |: |:::|:||::||| ||
Db    244 VMSKVMEMFQPSAVVLQCGADSLSGDRLGCFNLTIKGHAKCVEFIKTFNLPLLMLGGGGY 303

Qy    316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
          |||||||| ||||||:  |  |:||||:|||||||: ||: || | | || : :|:|:
Db    304 TIRNVARCWTYETAVALDSEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy    376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
          | | |  | ||| || |          |   ||          ||    | | || ||
Db    364 QRLFENLRMLPHAPGVQMQAVAE-----DSIHDD----------SG----EEDEDDPDKR 404

Qy    436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
          :   | |   ::|:    ::|  |   |
Db    405 I---SIR---SSDKRIACDEEFSDSEDE 426


RESULT   4
T20163
hypothetical protein C53A5.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T20163
R;Mortimore, B.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19232
A;Accession: T20163
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-461 <WIL>
A;Cross-references: EMBL:Z81486; PIDN:CAB03984.1; GSPDB:GN00023; CESP:C53A5.3
A;Experimental source: clone C53A5
C;Genetics:
A;Gene: CESP:C53A5.3
A;Map position: 5
```

```
A;Introns: 58/3; 98/1; 387/1
C;Superfamily: RPD3 protein; RPD3/acuC homology


  Query Match             53.4%;  Score 1359.5;  DB 2;  Length 461;
  Best Local Similarity   57.3%;  Pred. No. 2.4e-100;
  Matches  262;  Conservative   65;  Mismatches  109;  Indels   21;  Gaps    5;

Qy      12 SGP--DGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLA 69
           :|| :  |||||:|:|:  ||:|||||||| |||||||| | |:::| |:| ||| ||  |
Db       4 NGPLMEHGKRRVAYYYDSNIGNYYYGQGHVMKPHRIRMTHHLVLNYGLYRNLEIFRPFPA 63

Qy      70 DASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGS 129
             |: |||| ||: || | :|:::   |  : : :|||||||:|||||::||: |:||| 
Db      64 SFEDMTRFHSDEYMTFLKSANPDNL--KSFNKQMLKFNVGEDCPLFDGLYEFCQLSSGGS 121

Qy     130 IGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDV 189
           : || |||:|  |||||| |||||||||||||||| |||||||||||  |||||:|||| 
Db     122 LAAATKLNKQKVDIAINWMGGLHHAKKSEASGFCYTNDIVLGILELLKYHKRVLYVDIDV 181

Qy     190 HHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESF 249
           |||||||||||||||||||||||:|||||||| ::|:|| ||| |:::|||| ||: | |: 
Db     182 HHGDGVEEAFYTTDRVMTVSFHKYGDFFPGTGDLKDIGAGKGKLYSVNVPLRDGITDVSY 241

Qy     250 RSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMV 309
           :|:|:|:: |||| : | ||||||||||:|||| |||::||| :| || ||||||||: 
Db     242 QSIFKPIMTKVMERFDPCAVVLQCGADSLNGDRLGPFNLTLKGHGECARFFRSYNVPLMM 301

Qy     310 LGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPK 369
           :|| ||| ||||||| |||::||  |  |:||||:|||||||:| ||:: |   | |: 
Db     302 VGGGGYTPRNVARCWTYETSIAVDKEVPNELPYNDYFEYFGPNYRLHIESSNAANENSSD 361

Qy     370 DMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDS 429
            : :::  ::  |  |   |||| :  |     |    ||              :  :| 
Db     362 MLAKLQTDVIANLEQLTFVPSVQMRPIP--EDALSALNDD-------------SLIADQ 405

Qy     430 DDDDKPLHGYSCRGGATTDRD-STGEDEMDDDNPEPD 465
           : || |      |    | |   || | ||   | |
Db     406 ANPDKRLPPQITDGMIQDDGDFYDGEREGDDRRNESD 442


RESULT   5
S22284
transcription regulator RPD3 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein N0305; protein YNL330c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 23-Mar-2001
C;Accession: S22284; S51336; S51285; S55859; S57393; S63313; S63311
R;Vidal, M.; Gaber, R.F.
Mol. Cell. Biol. 11, 6317-6327, 1991
A;Title: RPD3 encodes a second factor required to achieve maximum positive and negative
A;Reference number: S22284; MUID:92049362
A;Accession: S22284
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-433 <VID>
A;Cross-references: GB:S66438; NID:g238961; PIDN:AAB20328.1; PID:g238962
R;van Dyck, L.; Pascual-Ahuir, A.; Goffeau, A.
submitted to the EMBL Data Library, December 1994
A;Description: A 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes
icated gene for a putative aryl-alcohol dehydrogenase.
A;Reference number: S51334
A;Accession: S51336
A;Molecule type: DNA
A;Residues: 1-433 <VAN>
A;Cross-references: EMBL:X83226; NID:g642335; PIDN:CAA58228.1; PID:g642338
R;Nicaud, J.J.
submitted to the EMBL Data Library, January 1995
A;Description: Sequence analysis of a 13.9 Kb fragment of yeast chromosome XIV identifie
A;Reference number: S51285
A;Accession: S51285
A;Molecule type: DNA
A;Residues: 1-201 <NIC>
A;Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86368.1; PID:g854536
R;Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
Yeast 11, 567-572, 1995
A;Title: Sequencing analysis of a 15.4 kb fragment of yeast chromosome XIV identifies
A;Reference number: S55859; MUID:95373280
A;Accession: S55859
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-201 <MAF>
A;Cross-references: EMBL:Z46259; NID:g633655; PIDN:CAA86368.1; PID:g854536
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
R;van Dyck, L.; Pascual-Ahuir, A.; Purnelle, B.; Goffeau, A.
Yeast 11, 987-991, 1995
A;Title: An 8.2 kb DNA segment from chromosome XIV carries the RPD3 and PAS8 genes as
d gene for a putative aryl-alcohol dehydrogenase.
A;Reference number: S57391; MUID:96021610
A;Accession: S57393
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-433 <VAW>
A;Cross-references: EMBL:X83226; NID:g642335; PIDN:CAA58228.1; PID:g642338
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994
R;Coster, F.; Jonniaux, J.L.; Goffeau, A.; Purnelle, B.; Van Dyck, L.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63151
A;Accession: S63313
A;Molecule type: DNA
A;Residues: 1-433 <COS>
A;Cross-references: EMBL:Z71606; NID:g1302450; PIDN:CAA96263.1; PID:g1302451; GSPDB:G
A;Experimental source: strain S288C
R;Maftahi, M.; Nicaud, J.M.; Levesque, H.; Gaillardin, C.
submitted to the Protein Sequence Database, April 1996
A;Reference number: S63287
A;Accession: S63311
A;Molecule type: DNA
A;Residues: 1-201 <MAW>
A;Cross-references: EMBL:Z71606; GSPDB:GN00014; MIPS:YNL330c
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:RPD3; SDI2; SDS6; MIPS:YNL330c
A;Cross-references: SGD:S0005274; MIPS:YNL330c
A;Map position: 14L
C;Function:
A;Description: transcription regulation
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: nucleus; transcription regulation
F;23-322/Domain: RPD3/acuC homology <RAH1>


  Query Match             52.2%;  Score 1328.5;  DB 1;  Length 433;
  Best Local Similarity   57.9%;  Pred. No. 6.4e-98;
  Matches  238;  Conservative   74;  Mismatches   94;  Indels    5;  Gaps    2;

Qy      18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
           ||||:|||:  :|:| || ||||||||||||||||::| |:::|| |   |   :: :| 
Db      18 KRRVAYFYDADVGNYAYGAGHPMKPHRIRMAHSLIMNYGLYKKMEIYRAKPATKQEMCQF 77

Qy      78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
           |: ||:|||: |:|:::      |   :||||:||||||||:::|  | |||:  | :|| 
Db      78 HTDEYIDFLSRVTPDNL--EMFKRESVKFNVGDDCPVFDGLYEYCSISGGGSMEGAARLN 135

Qy     138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEE 197
           |   |:|:|: |||||||||||||||||:|||||||:|||:   |||||||||||||||| 
Db     136 RGKCDVAVNYAGGLHHAKKSEASGFCYLNDIVLGIIELLRYHPRVLYIDIDVHHGDGVEE 195

Qy     198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
           |||||||||| |||||:|:|||||| :||:|    || ||:|||| ||:|| ::||:| |:| 
Db     196 AFYTTDRVMTCSFHKYGEFFPGTGELRDIGVGAGKNYAVNVPLRDGIDDATYRSVFEPVI 255

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
           :|:|| ||| ||||||| |||||||||||||||::|||:|: :::|: :|:||:|| |||: 
Db     256 KKIMEWYQPSAVVLQCGGDSLSGDRLGCFNLSMEGHANCVNYVKSFGIPMMVVGGGGYTM 315
```

```
Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNT 377
Db     316 RNVARTWCFETGLLNNVVLDKDLPYNEYYEYYGPDYKLSVRPSNMFNVNTPEYLDKVMTN 375

Qy     378 LLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESD 428
Db     376 IFANLENTKYAPSVQLNHTPRDAEDLGDVEED---SAEAKDTKGGSQYARD 423


RESULT   6
T23963
hypothetical protein R06C1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T23963; T24071
R;Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19824
A;Accession: T23963
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-465 <WIL>
A;Cross-references: EMBL:Z81106; PIDN:CAB03224.1; GSPDB:GN00019; CESP:R06C1.1
A;Experimental source: clone R06C1
R;Kershaw, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19837
A;Accession: T24071
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-465 <WI2>
A;Cross-references: EMBL:Z81108; PIDN:CAB03240.1; GSPDB:GN00019; CESP:R06C1.1
A;Experimental source: clone R09B3
C;Genetics:
A;Gene: CESP:R06C1.1
A;Map position: 1
A;Introns: 53/3; 212/1; 279/1; 326/1
C;Superfamily: RPD3 protein; RPD3/acuC homology


  Query Match             51.3%;  Score 1305.5;  DB 2;  Length 465;
  Best Local Similarity   61.9%;  Pred. No. 4.8e-96;
  Matches  234;  Conservative   67;  Mismatches   74;  Indels    3;  Gaps    2;

Qy      18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
Db       7 KSRVSYYYDGDFGNFYYGQGHPMKPHRVRMTHSLIVNYGLYRKLNVMRPARASFSEITRY 66

Qy      78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
Db      67 HSDDYINFLRNVKSDNMS--TFTDQMARFSVGEDCPVFDGMYEFCQLSCGGSLAAAARLN 124

Qy     138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEE 197
Db     125 RQESEIAINWMGGLHHAKKSEASGFCYSNDIVLAILELLKHHKRVLYIDIDVHHGDGVEE 184

Qy     198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
Db     185 AFYTTDRVMTVSFHKHGEYFPGTGDLKDVGAGSGKYYALNVPLRDGVDDVTYERIFRTIM 244

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
Db     245 GEVMARFQPEAVVLQCGADSLAGDRLGVFNLTTYGHGKCVEYMKSFNVPLLLVGGGGYTI 304

Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDP-SPMENLNTPKDMERIRN 376
Db     305 RNVSRCWLYETAIALNQEVSDDLPLHDYFDYFIPDYKLHIKPLAALSNFNTPEFIDQTIV 364

Qy     377 TLLEQLSGLIHAPSVQFQ 394
Db     365 ALLENLKQLPHVPSVQMQ 382
```

```
RESULT   7
JC7102
histone deacetylase (EC 3.5.1.-) - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 11-May-2000
C;Accession: JC7102
R;Mahlknecht, U.; Hoelzer, D.; Bucala, R.; Verdin, E.
Biochem. Biophys. Res. Commun. 263, 482-490, 1999
A;Title: Cloning and characterization of the murine histone deacetylase (HDAC3).
A;Reference number: JC7102; MUID:99423490
A;Accession: JC7102
A;Molecule type: mRNA
A;Residues: 1-424 <MAH>
A;Cross-references: GB:AF074881
C;Genetics:
A;Gene: hdac3
A;Map position: 18B3
A;Introns: 19/1; 46/3; 94/2; 121/3; 140/3; 159/2; 204/1; 231/1; 252/3; 277/2; 307/2;
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase; phosphoprotein
F;7-307/Domain: RPD3/acuC homology <RAH1>
F;232,282,404/Binding site: phosphate (Tyr) (covalent) #status predicted


  Query Match             50.8%;  Score 1292.5;  DB 2;  Length 424;
  Best Local Similarity   55.8%;  Pred. No. 4.6e-95;
  Matches  235;  Conservative   74;  Mismatches  103;  Indels    9;  Gaps    5;

Qy      19 RRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFH 78
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      79 SPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNR 138
Db      63 SEDYIDFLQRVSPTNM--QGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNN 120

Qy     139 QDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEA 198
Db     121 KICDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEA 180

Qy     199 FYTTDRVMTVSFHKFGD-FFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
Db     181 FYLTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVI 240

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
Db     241 SQVVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTV 300

Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP-MENLNTPKDMERIRN 376
Db     301 RNVARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQIRQ 360

Qy     377 TLLEQLSGLIHAPSVQFQHTPPVNRVLDEPED-DMETRPKPRXWSGTAT----YESDSDD 431
Db     361 TIFENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDN 420

Qy     432 D 432
Db     421 D 421


RESULT   8
JC5834
histone deacetylase (EC 3.5.1.-) 3 - human
C;Species: Homo sapiens (man)
C;Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 12-Nov-1999
C;Accession: JC5834
R;Dangond, F.; Hafler, D.A.; Tong, J.K.; Randall, J.; Kojima, R.; Utku, N.; Gullans,
Biochem. Biophys. Res. Commun. 242, 648-652, 1998
A;Title: Differential display cloning of a novel human histone deacetylase (HDAC3) cD
A;Reference number: JC5834; MUID:98125547
A;Accession: JC5834
A;Status: nucleic acid sequence not shown
```

A;Molecule type: mRNA
A;Residues: 1-428 <DAN>
A;Cross-references: GB:U66914; NID:g2934900; PIDN:AAC52038.1; PID:g2326173
C;Comment: This enzyme plays a role in multiple and complex cellular pathways of immune
gets for the treatment of cancer and autoimmunity.
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase; phosphoprotein
F;7-307/Domain: RPD3/acuC homology <RAH1>
F;232,282,404/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 50.5%; Score 1285.5; DB 2; Length 428;
Best Local Similarity 55.6%; Pred. No. 1.7e-94;
Matches 234; Conservative 74; Mismatches 104; Indels 9; Gaps 5;

```
Qy      19 RRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFH 78
           : |:|||:| :|:::|| |||||||: : |||::|| |:::: : :|  |   |: |||
Db       3 KTVAYFYDPDVGNFHYGAGHPMKPHRLALTHSLVLHYGLYKKMIVFKPYQASQHDMCRFH 62

Qy      79 SPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNR 138
           | :|:|||  ||| :|     ::|  ||||:||||| |||:||    | |:  | :||
Db      63 SEDYIDFLQRVSPTNM--QGFTKSLNAFNVGDDCPVFPGLFEFCSRYTGASLQGATQLNN 120

Qy     139 QDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEA 198
           :  |||||| |||||||| |||||||||||||:|||||||   |||||||||:||||||:||
Db     121 KICDIAINWAGGLHHAKKFEASGFCYVNDIVIGILELLKYHPRVLYIDIDIHHGDGVQEA 180

Qy     199 FYTTDRVMTVSFHKFGD-FFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
           || |||||||||||:|: |||||| : :|||| |:|| ||||| ||:||:|:: ||:|:|
Db     181 FYLTDRVMTVSFHKYGNYFFPGTGDMYEVGAESGRYYCLNVPLRDGIDDQSYKHLFQPVI 240

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
            :|:: |||  :|||||||||  ||||||||||:|| :|: :::|:|:|:||:|||| |||:
Db     241 NQVVDFYQPTCIVLQCGADSLGCDRLGCFNLSIRGHGECVEYVKSFNIPLLVLGGGGYTV 300

Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP-MENLNTPKDMERIRN 376
           |||||||  |||:: |      :|||:|||||| ||:||| | |  :|| |: : :::|
Db     301 RNVARCWTYETSLLVEEAISEELPYSEYFEYFAPDFTLHPDVSTRIENQNSRQYLDQILQ 360

Qy     377 TLLEQLSGLIHAPSVQFQHTPPVNRVLDEPED-DMETRPKPRXWSGTAT----YESDSDD 431
           |: | |  | ||||||    |      |  :: | | |     :|       |: | |:
Db     361 TIFENLKMLNHAPSVQIHDVPADLLTYDRTDEADAEERGPEENYSRPEAPNEFYDGDHDN 420

Qy     432 D 432
           |
Db     421 D 421
```

RESULT 9
T40300
histone deacetylase (EC 3.5.1.-) - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: cryptic loci regulator protein clr6
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T40300; T43796
R;Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duesterh
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21919
A;Accession: T40300
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-405 <LYN>
A;Cross-references: EMBL:AL023589; PIDN:CAA19053.1; GSPDB:GN00067; SPDB:SPBC36.05c
A;Experimental source: strain 972h-; cosmid c36
R;Grewal, S.I.; Bonaduce, M.J.; Klar, A.J.
Genetics 150, 563-576, 1998
A;Title: Histone deacetylase homologs regulate epigenetic inheritance of transcriptional
A;Reference number: Z22685; MUID:98429513
A;Accession: T43796
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-405 <GRE>
A;Cross-references: EMBL:AF064206; NID:g4159996; PIDN:AAD05211.1; PID:g4159997
A;Experimental source: strain 972
C;Genetics:
A;Gene: clr6; SPBC36.05c
A;Map position: 2
A;Introns: 14/1; 34/3; 51/3; 91/1; 101/2; 116/1; 131/2; 209/3; 240/3; 277/1; 324/1; 3
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase; transcription regulation

Query Match 50.1%; Score 1275; DB 2; Length 405;
Best Local Similarity 59.5%; Pred. No. 1.1e-93;
Matches 226; Conservative 70; Mismatches 82; Indels 2; Gaps 1;

```
Qy      18 KRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRF 77
           |::|||||:  :|:|:||  |||||||:|| |:|:::|:|: :| :  |  |  :|: |
Db       5 KKKVSYFYDEDVGNYHYGPQHPMKPHRVRMVHNLVVNYNLYEKLNVITPVRATRNDMTRC 64

Qy      78 HSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLN 137
           |: ||::||  |:|::|       :  :||||:|||||||||::|| ||||||||| :||
Db      65 HTDEYIEFLWRVTPDTM--EKFQPHQLKFNVGDDCPVFDGLYEFCSISAGGSIGAAQELN 122

Qy     138 RQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEE 197
             :|:|||| |||||||| |||||||||| |  |||||  :|||||||||||||||||
Db     123 SGNAEIAINWAGGLHHAKKREASGFCYVNDIALAALELLKYHQRVLYIDIDVHHGDGVEE 182

Qy     198 AFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLI 257
           |||||||| |||||::|||||||:| |   || ||:|||| ||:|||: |:|:|:|
Db     183 FFYTTDRVMTCSFHKFGEYFPGTGHIKDTGIGTGKNYAVNVPLRDGIDDESYESVFKPVI 242

Qy     258 QKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTI 317
            :|: ::||||:|||| |||:|||||||||||:|||: |: |::|:|:|:|:: :|| |||:
Db     243 SHIMQWFRPEAVILQCGTDSLAGDRLGCFNLSMKGHSMCVDFVKSFNLPMICVGGGGYTV 302

Qy     318 RNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNT 377
           ||||| | ||| :  | | |  ||||:|:|:|||| |:|  : ||| || : :: | :
Db     303 RNVARVWTYETGLLAGEELDENLPYNDYLQYYGPDYKLNVLSNNMENHNTRQYLDSITSE 362

Qy     378 LLEQLSGLIHAPSVQFQHTP 397
           ::| |  |  |||||   ||
Db     363 IIENLRNLSFAPSVQMHKTP 382
```

RESULT 10
T47443
hypothetical protein T18B22.80 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
C;Accession: T47443
R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Lemcke,
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24467
A;Accession: T47443
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-419 <JOR>
A;Cross-references: EMBL:AL138652
A;Experimental source: cultivar Columbia; BAC clone T18B22
C;Genetics:
A;Map position: 3
A;Introns: 47/3; 87/1; 112/1; 185/2; 206/3; 237/3; 270/1; 286/3; 317/1; 340/3; 373/3;
A;Note: T18B22.80
C;Superfamily: RPD3 protein; RPD3/acuC homology

Query Match 47.4%; Score 1207; DB 2; Length 419;
Best Local Similarity 53.4%; Pred. No. 2.8e-88;
Matches 225; Conservative 71; Mismatches 109; Indels 16; Gaps 5;

```
Qy      21 VSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHSP 80
           | ::    :|  |:|  |||||||: | | ||: | || ::|: ||  |   :: :||||
Db       4 VFFWVAGDVGSVYFGPNHPMKPHRLCMTHHLILAYGLHSKMEVYRPHKAYPIEMAQFHSP 63

Qy      81 EYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGS1GAAVKLNRQD 140
```

```
          :|||:||  ::|||:         : |:|:|||||||: ||:||:  |||:| || :|| :
Db     64 DYVEFLQRINPENQN--LFPNEMARYNLGEDCPVFEDLFEFCQLYAGGTIDAARRLNNKL 121

Qy    141 ADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFY 200
          |||||| |||||||| :|||||:||:|||||||||   ||||||||||||||||||||
Db    122 CDIAINWAGGLHHAKKCDASGFCYINDLVLGILELLKHHPRVLYIDIDVHHGDGVEEAFY 181

Qy    201 TTDRVMTVSFHKFGD-FFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQK 259
          |||||||||||||| |||||| :::::|  :|||:||:|||| ||:|| ||  |||
Db    182 FTDRVMTVSFHKFGDKFFPGTGDVKEIGEREGKFYAINVPLKDGIDDSSFNRLFR----T 237

Qy    260 VMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTIRN 319
          |:|:||| }:|||||||||: |||||||||: |||:|::|:: :|:||:| || |||  |
Db    238 VVEIYQPGAIVLQCGADSLARDRLGCFNLSIDGHAECVKFVKKFNLPLLVTGGGGYTKEN 297

Qy    320 VARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNTLL 379
          |||||  || : :  | |::| |:| :|| ||::| :    :|||||   :  |:  :|
Db    298 VARCWTVETGILLDTELPNEIPENDYIKYFAPDFSLKIPGGHIENLNTKSYISSIKVQIL 357

Qy    380 EQLSGLIHAPSVQFQHTPPVNRVLDEPED--------DMETRPKPRXWSGTATYESDSDD 431
          | |  : ||||||  |  ||   : |  ||        |  :| | :       :: |:|:
Db    358 ENLRYIQHAPSVQMQEVPPDFYIPDFDEDEQNPDVRADQRSRDK-QIQRDDEYFDGDNDN 416

Qy    432 D 432
          |
Db    417 D 417


RESULT  11
T11643
histone deacetylase (EC 3.5.1.-) - fission yeast (Schizosaccharomyces pombe)
N;Alternate names: Phd1 protein
C;Species: Schizosaccharomyces pombe
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T11643; T43300
R;Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, December 1997
A;Reference number: Z17304
A;Accession: T11643
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-434 <WOO>
A;Cross-references: EMBL:AL021046
R;Kim, Y.B.; Honda, A.; Yoshida, M.; Horinouchi, S.
FEBS Lett. 436, 193-196, 1998
A;Title: Phd1+, a histone deacetylase gene of Schizosaccharomyces pombe, is required for
A;Reference number: Z22404; MUID:98452937
A;Accession: T43300
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-434 <KIM>
A;Cross-references: EMBL:AB008888; NID:g2641698; PIDN:BAA23598.1; PID:g2641699
C;Genetics:
A;Gene: hda1; Phd1; SPAC3G9.07c
A;Map position: 1L
C;Function:
A;Description: involved in the meiotic cell cycle
C;Superfamily: RPD3 protein; RPD3/acuC homology
C;Keywords: hydrolase
F;30-330/Domain: RPD3/acuC homology <RAC>


  Query Match             44.9%;  Score 1143.5;  DB 2;  Length 434;
  Best Local Similarity   53.8%;  Pred. No. 3.3e-83;
  Matches  207;  Conservative   73;  Mismatches  100;  Indels    5;  Gaps    4;

Qy     17 RKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGR 76
          :|:||:|  :  :|:|:||  |||||||| : : |:: | || :: :  | :|   ::
Db     24 QKKRVTYHLDEQVGNYHYGDKHPMKPHRITITNHLVMGYGLHNKMSVFSPRMATFGEMSE 83

Qy     77 FHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKL 136
          ||  :|:|||  |:|::      |   ::||:|:|||||| ::| : ||| |: |: ||
Db     84 FHREDYLDFLKRVTPDNA--EQFADKFQQFNIGDDCPVFDGTYEFSQRSAGASLDASRKL 141

Qy    137 NRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVE 196
          :   |||||| |||||||: |||||||||||| || :|: | ||||||||:||||||:
Db    142 VQGQTDIAINWSGGLHHAKRGEASGFCYVNDIVLAILNMLRFFPRVLYIDIDIHHGDGVQ 201

Qy    197 EAFYTTDRVMTVSFHKF-GDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
          :||| :|||:|||||: ||||| ||:  : | : |||:|||||| ||: || : |||:
Db    202 QAFYESDRVLTVSFHKYNGDFFPATGNFDENGVKGGKYFALNVPLEDGIGDEQYTSLFKS 261

Qy    256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
          :|:  :  :|| |:|||||||||  |||| ||||:  | :|:|| ||:|:|:::|:|| ||
Db    262 IIEPTINTFQPSAIVLQCGADSLGYDRLGVFNLSIHAHGECVRFTRSFNIPMLVVGGGGY 321

Qy    316 TIRNVARCWCYETAVAVGVEPDNKLPYNE-YFEYFGPDYTLHVD-PSPMENLNTPKDMER 373
          |:|||||  ||||||:: |  :  ::||    |:|:| ||||||    : :|| |||| :|
Db    322 TLRNVARAWCYETSICVNEQIPSELPRETLYYEFFAPDYTLHPRLTTKIENKNTPKALED 381

Qy    374 IRNTLLEQLSGLIHAPSVQFQHTPP 398
          :|   ||||  |  ||||| |  ||
Db    382 LRIRALEQLRYLGGAPSVQMQQIPP 406


RESULT  12
T19067
hypothetical protein C08B11.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C;Accession: T19067
R;Sims, M.
submitted to the EMBL Data Library, November 1994
A;Reference number: Z19069
A;Accession: T19067
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-507 <WIL>
A;Cross-references: EMBL:Z46676; PIDN:CAA86662.1; GSPDB:GN00020; CESP:C08B11.2
A;Experimental source: clone C08B11
C;Genetics:
A;Gene: CESP:C08B11.2
A;Map position: 2
A;Introns: 45/1; 298/2
C;Superfamily: RPD3 protein; RPD3/acuC homology


  Query Match             40.9%;  Score 1041.5;  DB 2;  Length 507;
  Best Local Similarity   44.6%;  Pred. No. 5.3e-75;
  Matches  221;  Conservative   70;  Mismatches  171;  Indels   33;  Gaps    7;

Qy      4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
          |   |  | | | :|| |:|:|   :| ::||| ||||| |: : : |:: | : : : :
Db     14 DNDEIIEPDGADVKKRNVAYYYHKDVGHFHYGQLHPMKPQRLVVCNDLVVSYEMPKYMTV 73

Qy     64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
                ||:||  ||: :||:|| :|:|: :|       ||:||:|||||:| ||:|:|
Db     74 VESPKLDAADISVFHTEDYVNFLQTVTPK-LGLTMPDDVLRQFNIGEDCPIFAGLWDYCT 132

Qy    124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
            ||||:  | :|| :  || ||| ||||||||||||||||||||||||||||  ||||
Db    133 LYAGGSVEGARRLNHKMNDIVINWPGGLHHAKKSEASGFCYVNDIVLGILELLKYHKRVL 192

Qy    184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
          |||||:||||||:|||  :|||||||||:|| :|||:| | | |   |||:|:||||
Db    193 YIDIDIHHGDGVQEAFNNSDRVMTVSFHRFGQYFPGSGSIMDKGVGPGKYFAINVPLMAA 252

Qy    244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
          : || :  ||  :|  | | : |||:|||||:|||  |||| | ||   ||  :::::|
Db    253 IRDEPYLKLFESVISGVEENFNPEAIVLQCGSDSLCEDRLGQFALSFNAHARAVKYVKSL 312

Qy    304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNE-YFEYFGPDYTLHVDPSPM 362
          |||||| |||:|||||||  || | :|:  |:::|    |  || |           |
Db    313 GKPLMVLGGGGYTLRNVARCWALETGVILGLRMDDEIPGTSLYSHYFTPRLLRPNLVPKM 372

Qy    363 ENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGT 422
```

```
Db    373 NDANSAAYLASIEKETLACLRMIRGAPSVQMQNIVGIR--LDEIEQIEENERLQKSSKSS 430

Qy    423 ATYESDS----------DDDDKP-------------LHGYSCRGGATTDR---DSTGE 454
Db    431 IEYEVGKVSEKMEEECFVEEDSKPPSFPPGQDPRRIGQYWGYD-RSGLAPPRSHSDVIEE 489

Qy    455 DEMDDDNPEPDVNPP 469
Db    490 AKYEDRDRRKDLNIP 504
```

RESULT 13
S64211
probable transcription regulator YGL194c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G1330
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
C;Accession: S64211; S62051
R;Bruschi, C.V.; Coglievina, M.; Bertani, I.; Klima, R.; Zaccaria, P.; Delneri, D.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64183
A;Accession: S64211
A;Molecule type: DNA
A;Residues: 1-452 <BRU>
A;Cross-references: EMBL:Z72716; NID:g1322818; PIDN:CAA96906.1; PID:g1322819; MIPS:YGL19
A;Experimental source: strain S288C
R;Coglievina, M.; Delneri, D.; Zaccaria, P.; Klima, R.; Bertani, I.; Bruschi, C.V.
submitted to the EMBL Data Library, September 1995
A;Description: A 6.7 Kb fragment from chromosome VII of Saccharomyces cerevisiae contain
A;Reference number: S62051
A;Accession: S62051
A;Molecule type: DNA
A;Residues: 1-351,'HSGTHS',358,'RIIHFI',365-452 <COG>
A;Cross-references: EMBL:X91837; NID:g1177627; PIDN:CAA62950.1; PID:g1177634
A;Experimental source: strain FY1679
C;Genetics:
A;Gene: SGD:HOS2
A;Cross-references: SGD:S0003162; MIPS:YGL194c
A;Map position: 7L
C;Superfamily: RPD3 protein; RPD3/acuC homology
F;30-331/Domain: RPD3/acuC homology <RAH1>

  Query Match 39.4%; Score 1003.5; DB 2; Length 452;
  Best Local Similarity 44.0%; Pred. No. 4.8e-72;
  Matches 198; Conservative 86; Mismatches 127; Indels 39; Gaps 9;

```
Qy     20 RVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHS 79
Db     27 RVSYHFNSKVSHYHYGVKHPMKPFRLMLTDHLVSSYGLHKIMDLYETRSATRDELLQFHS 86

Qy     80 PEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQ 139
Db     87 EDYVNFLSKVSPEN-ANKLPRGTLENFNIGDDCPIFQNLYDYTTLYTGASLDATRKLINN 145

Qy    140 DADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAF 199
Db    146 QSDIAINWSGGLHHAKKNSPSGFCYVNDIVLSILNLLRYHPRILYIDIDLHHGDGVQEAF 205

Qy    200 YTTDRVMTVSFHKF-GDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQ 258
Db    206 YTTDRVFTLSFHKYNGEFFPGTGDLTEIGCDKGKHFALNVPLEDGIDDDSYINLFKSIVD 265

Qy    259 KVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTIR 318
Db    266 PLIMTFKPTLIVQQCGADSLGHDRLGCFNLNIKAHGECVKFVKSFGLPMLVVGGGGYTPR 325

Qy    319 NVARCWCYETAVAVGVEPDNKLPYNEYF-EYFGPDYTLHVDPSPM-----ENLNTPKDME 372
Db    326 NVSRLWTYETGILNDVLLPEDIPEDIPFRDSFGPDYSLY----PMLDDLYENKNSKKLLE 381

Qy    373 RIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDD 432
```

```
Db    382 DIRIRCLENIRYLQGAPSVRM--------------DAECIP-------TQDISALTEEE 419

Qy    433 DKPLHGYSCRGGATTDRDSTGE-DEMDDDN 461
Db    420 DKIIQEMN----EETEADSSNRLEEMEKEN 445
```

RESULT 14
D70388
acetoin utilization protein - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 31-Mar-2000
C;Accession: D70388
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666
A;Accession: D70388
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-375 <AQF>
A;Cross-references: GB:AE000719; NID:g2983517; PIDN:AAC07100.1; PID:g2983524; GB:AE00
A;Experimental source: strain VF5
C;Genetics:
A;Gene: acuC1
C;Superfamily: acetylpolyamine aminohydrolase; RPD3/acuC homology
F;6-306/Domain: RPD3/acuC homology <RAH1>

  Query Match 18.0%; Score 457.5; DB 2; Length 375;
  Best Local Similarity 34.0%; Pred. No. 9.7e-29;
  Matches 127; Conservative 65; Mismatches 140; Indels 41; Gaps 14;

```
Qy     30 GDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPEYVDFLASV 89
Db     13 GKYRYPKNHPLKIPRVSLLLRFLDAMNLIDEKELIKSRPATKEELLLFHTEDYINTLME- 71

Qy     90 SPESMGDPSAARNLRRFNV-GEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDADIAINWG 148
Db     72 AERCQCVPKGAR--EKYNIGGYENPVSYAMFTGSSLATGSTVQAIEEFLK--GNVAFNPA 127

Qy    149 GGLHHAKKSEASGFCYVNDIVLGILELLKM-FKRVLYIDIDVHHGDGVEEAFYTTDRVMT 207
Db    128 GGMHHAFKSRANGFCYINDPAVGIEYLRKKGFKRILYIDLDAHHCDGVQEAFYDTDQVFV 187

Qy    208 VSFHKFGDF-FP-GTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQ 265
Db    188 LSLHQSPEYAFPFEKGFLEEIGEGKGKGYNLNIPLPKGLNDNEFLFALEKSLEIVKEVFE 247

Qy    266 PEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMV------LGGEGYTIRN 319
Db    248 PEVYLLQLGTDPLLEDYLSKFNLS------NVAFLKAFNIVREVFGEGVYLGGGGYHPYA 301

Qy    320 VARC----WCYETAVAVGVEPDNK---LPYNEYFEYFGPDYTLHVDPSPMENLNTPKD-- 370
Db    302 LARAWTLIWCELSGREVPEKLNNKAKELLKSIDFEEFDDE----VDRSYM--LETLKDPW 355

Qy    371 -----MERIRNTL 378
Db    356 RGGEVRKEVKDTL 368
```

RESULT 15
S54089
hypothetical protein YPR068c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein YP9499.23c
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 06-Feb-1998
C;Accession: S54089
R;Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995

```
A;Reference number: S54059
A;Accession: S54089
A;Molecule type: DNA
A;Residues: 1-470 <BAD>
A;Cross-references: EMBL:Z49219; NID:g805025; PID:g805048; MIPS:YPR068c
A;Experimental source: strain AB972
C;Genetics:
A;Gene: SGD:HOS1
A;Cross-references: SGD:S0006272; MIPS:YPR068c
A;Map position: 16R
C;Superfamily: RPD3/acuC homology
F;9-383/Domain: RPD3/acuC homology <RAH1>


  Query Match            17.1%;  Score 436;  DB 2;  Length 470;
  Best Local Similarity   28.0%;  Pred. No. 6.8e-27;
  Matches  104;  Conservative   67;  Mismatches  116;  Indels   84;  Gaps   10;

Qy      39 PMKPH-RIRMAHSLIIHYHLHRRL-EISRPSLADASDIGRFHSPEYVDFLAS-----VSP 91
           |   | : :: :|||  | | :   |:     |   |:  ||| |:|:| :     :
Db      20 PCNNHQKSQLTYSLINAYDLLQHFDEVLTFPYARKDDLLEFHSKSYIDYLINGRFNKMMA 79

Qy      92 ESMGDP----------------------SAARNLRRF---------------------- 106
           : : :|                      : :::|:||
Db      80 QDVNNPMVESKWSELSELADNWNEKIDYNPSQDLQRFTTRENLYNYYLNHSQALENNMDC 139

Qy     107 ------------------------NVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDA 141
                                   |:  |||:|  |  :|:   | ::     |:  :
Db     140 INNSEVPTNDKPTDTYILNSETKQYNLEGDCPIFSYLPMYCQVITGATLNLLDHLSPTER 199

Qy     142 DIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLK-MFKRVLYIDIDVHHGDGVEEAFY 200
           | ||| || ||| |  ||||||:||:|| |  | |    :: |:| |:|||||||:||
Db     200 LIGINWDGGRHHAFKQRASGFCYINDVVLLIQRLRKAKLNKITYVDFDLHHGDGVEKAFQ 259

Qy     201 TTDRVMTVSFHKF-GDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQK 259
           : :: |:| | :   ||||||  : |   :|     :|:||  | ||     :   ::
Db     260 YSKQIQTISVHLYEPGFFPGTGSLSDSRKDKN---VVNIPLKHGCDDNYLELIASKIVNP 316

Qy     260 VMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHAD-CLRFLRSY-NVPLMVLGGEGYTI 317
           ::| ::|||::::|| | | |||   : |:::| :   :  ::||    : :||| ||
Db     317 LIERHEPEALIIECGGDGLLGDRFNEWQLTIRGLSRIIINIMKSYPRAHIFLLGGGGYND 376

Qy     318 RNVARCWCYET 328
           ::| : | |
Db     377 LLMSRFYTYLT 387


Search completed: April 26, 2002, 17:52:47
Job time: 110 sec
```

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:51:24 ; Search time 21.38 Seconds
(without alignments)
495.746 Million cell updates/sec

Title: US-09-645-337-4
Perfect score: 2544
Sequence: 1 MEADESGISLPSGPDGRKRR.....TGEDEMDDDNPEPDVNPPSS 471

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1734	68.2	458	4	US-09-282-305-2	Sequence 2, Appli
2	1551.5	61.0	517	4	US-09-282-305-8	Sequence 8, Appli
3	1520	59.7	439	4	US-09-282-305-6	Sequence 6, Appli
4	1448.5	56.9	482	1	US-08-528-255A-1	Sequence 1, Appli
5	1448.5	56.9	482	1	US-08-717-365-1	Sequence 1, Appli
6	1280.5	50.3	351	4	US-09-282-305-4	Sequence 4, Appli
7	1249.5	49.1	432	4	US-09-282-305-10	Sequence 10, Appl
8	247.5	9.7	335	4	US-09-446-504-80	Sequence 80, Appl
9	104.5	4.1	6095	4	US-09-144-085-2	Sequence 2, Appli
10	101.5	4.0	4150	4	US-09-428-517-2	Sequence 2, Appli
11	93	3.7	593	1	US-08-296-362-2	Sequence 2, Appli
12	91	3.6	396	4	US-08-867-611-2	Sequence 2, Appli
13	91	3.6	396	5	PCT-US92-06965A-7	Sequence 7, Appli
14	88	3.5	820	1	US-08-291-896-2	Sequence 2, Appli
15	88	3.5	820	2	US-08-485-278-2	Sequence 2, Appli
16	87.5	3.4	473	4	US-09-286-691-15	Sequence 15, Appl
17	87.5	3.4	473	4	US-09-687-147-15	Sequence 15, Appl
18	87.5	3.4	1135	2	US-08-469-537A-97	Sequence 97, Appl
19	86	3.4	4928	4	US-09-036-987A-5	Sequence 5, Appli
20	86	3.4	4928	4	US-09-370-700-5	Sequence 5, Appli
21	85.5	3.4	767	1	US-08-446-777-6	Sequence 6, Appli
22	85	3.3	438	3	US-08-486-099-105	Sequence 105, App
23	85	3.3	438	3	US-08-360-107A-115	Sequence 115, App
24	85	3.3	438	3	US-08-484-223B-105	Sequence 105, App
25	85	3.3	438	3	US-08-919-597-105	Sequence 105, App
26	85	3.3	438	3	US-08-475-668A-105	Sequence 105, App
27	85	3.3	438	3	US-08-485-551A-105	Sequence 105, App
28	85	3.3	438	3	US-08-471-913A-105	Sequence 105, App
29	85	3.3	438	4	US-08-485-264A-105	Sequence 105, App
30	85	3.3	550	1	US-08-279-700-18	Sequence 18, Appl
31	85	3.3	550	1	US-08-279-700-20	Sequence 20, Appl
32	85	3.3	550	1	US-08-279-700-22	Sequence 22, Appl
33	85	3.3	550	4	US-09-230-944-20	Sequence 20, Appl
34	84.5	3.3	767	1	US-08-446-777-8	Sequence 8, Appli
35	84	3.3	541	4	US-08-867-611-8	Sequence 8, Appli
36	84	3.3	541	5	PCT-US92-06965A-13	Sequence 13, Appl
37	83.5	3.3	435	1	US-08-484-105-12	Sequence 12, Appl
38	83.5	3.3	435	1	US-08-484-106-12	Sequence 12, Appl
39	83	3.3	417	4	US-08-867-611-20	Sequence 20, Appl
40	83	3.3	417	5	PCT-US92-06965A-25	Sequence 25, Appl
41	83	3.3	550	1	US-08-348-891A-5	Sequence 5, Appli
42	83	3.3	550	2	US-08-905-817-5	Sequence 5, Appli
43	83	3.3	933	3	US-08-293-728-2	Sequence 2, Appli
44	83	3.3	933	4	US-09-421-868-2	Sequence 2, Appli
45	83	3.3	3567	2	US-07-642-734C-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-282-305-2
; Sequence 2, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; CURRENT FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Zea mays
US-09-282-305-2

Query Match 68.2%; Score 1734; DB 4; Length 458;
Best Local Similarity 69.3%; Pred. No. 1.5e-183;
Matches 328; Conservative 50; Mismatches 73; Indels 22; Gaps 7;

```
Qy       1 MEADESGISLPS---GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHL 57
           | |   | ||||   | |  :||||||||:||||||||||||||||||||||::|| |
Db       1 MAASGEGASLPSPAGGEDAHRRRVSYFYEPSIGDYYYGQGHPMKPHRIRMAHSLVVHYGL 60

Qy      58 HRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSA--ARNLRRFNVGEDCPVF 115
           || ||:|||  |  :|| |||| :|| |||| :    |:|     | ::||||||||||
Db      61 HRLLELSRPYPASEADIRRFHSDDYVAFLASAT----GNPGVLDPRAIKRFNVGEDCPVF 116

Qy     116 DGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILEL 175
           |||| ||:|||||||||||||||| |||| :|| ||||||||||||||||||||| ||||
Db     117 DGLFPFCQASAGGSIGAAVKLNRGDADITVNWAGGLHHAKKSEASGFCYVNDIVLAILEL 176

Qy     176 LKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYA 235
           ||  :|||||:|||||||||||||:||:|||||||||:|||||||||| |||| :||:||
Db     177 LKFHRRVLYVDIDVHHGDGVEEAFFTTNRVMTVSFHKYGDFFPGTGHITDVGAAEGKHYA 236

Qy     236 LNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHAD 295
           |||||:||:|| :|| ||: :|:|||||||: ||||||||||:|||||||||||||||||
Db     237 LNVPLSDGIDDTTFRGLFQCIIKKVMEVYQPDVVVLQCGADSLAGDRLGCFNLSVKGHAD 296

Qy     296 CLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTL 355
           ||||||||||:|||| ||||||||||||||||||||||||||||:|:|||||||||
Db     297 CLRFLRSYNVPMMVLGGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNDYYEYFGPDYTL 356
```

```
Qy    356 HVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRP- 414
          |: |  :|||||  ||:| |:| :|| || : | || ||   |      :| |:||: ||
Db    357 HIQPKSVENLNTTKDLENIKNMILENLSKIEHVPSTQFHDRPSDPEAPEEKEEDMDKRPP 416

Qy    415 -KPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDV 466
           : | ||| | |:||::| |           :  :| |   :| |: |: |:
Db    417 QRSRLWSGGA-YDSDTEDPDSL---------KSEGKDVTANFQMKDE-PKDDL 458


RESULT   2
US-09-282-305-8
; Sequence 8, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 8
;   LENGTH: 517
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-8


  Query Match           61.0%;  Score 1551.5;  DB 4;  Length 517;
  Best Local Similarity   66.4%;  Pred. No. 3.2e-163;
  Matches  288;  Conservative   59;  Mismatches   72;  Indels   15;  Gaps    5;

Qy      3 ADESGISLPS-GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
          |   | |||| ||||:|||| |||:| :|:|||||||||||||||| |||:  | |  ::
Db      6 AGSGGNSLPSVGPDGQKRRVCYFYDPDVGNYYYGQGHPMKPHRIRMTHSLLARYGLLNQM 65

Qy     62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDF 121
          :: ||: |   |: ||||: :|::|| ||:||:  |    | |:|||||||||||||: |
Db     66 QVYRPNPARDRDLCRFHADDYINFLRSVTPETQQD--QIRLLKRFNVGEDCPVFDGLYSF 123

Qy    122 CRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKR 181
          |:  || |:| |||||    |||||| |||||||| ||||||||||||| ||||||  :|
Db    124 CQTYAGASVGGAVKLN-HGHDIAINWSGGLHHAKKCEASGFCYVNDIVLAILELLKHHER 182

Qy    182 VLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLN 241
          |||:|||:|||||||||||||||||||||||||||:||||| |||:|  |||||:|||||:
Db    183 VLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDIGHSKGKYYSLNVPLD 242

Qy    242 DGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLR 301
          ||:|||::|||:|:: |||||::| ||||||||||||||||||||:||||:|:|::|
Db    243 DGIDDESYQSLFKPIMGKVMEVFRPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVRYMR 302

Qy    302 SYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP 361
          |:||||::||| |||||||||||||| ||:| ||::|:| |||:||||||||| ||
Db    303 SFNVPLLLLGGGGYTIRNVARCWCYETGVALGQEPEDKMPVNEYYEYFGPDYTLHVAPSN 362

Qy    362 MENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXW 419
          ||| || : :: ||: ||: || | ||||| ||   |   :   || :|| : |  |
Db    363 MENKNTRQQLDDIRSKLLDNLSKLRHAPSVHFQERVPDTEIPEQDEDQDDPDERHDP--- 419

Qy    420 SGTATYESDSDDDD 433
                :|| : ||
Db    420 ------DSDMEVDD 427


RESULT   3
US-09-282-305-6
; Sequence 6, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 439
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-6


  Query Match           59.7%;  Score 1520;  DB 4;  Length 439;
  Best Local Similarity   66.0%;  Pred. No. 7.5e-160;
  Matches  289;  Conservative   56;  Mismatches   77;  Indels   16;  Gaps    6;

Qy      3 ADESGISLP-SGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
          |   | ||| :| || |||| |||:  :|:||||||||||||||||| |:|:  | |  ::
Db      5 AGGGGNSLPTTGADGSKRRVCYFYDAEVGNYYYGQGHPMKPHRIRMTHALLGRYGLLDQM 64

Qy     62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDF 121
          :: ||  |   |: |||: :|| || ||:||:  |    | |:|||||||||||||: |
Db     65 QVFRPHPARDRDLCRFHADDYVSFLRSVTPETQQD--QIRALKRFNVGEDCPVFDGLYSF 122

Qy    122 CRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKR 181
          |:  |||||:| |||||    |||||| |||||||| |||||||||||| ||||||  :|
Db    123 CQTYAGGSVGGAVKLN-HGHDIAINWAGGLHHAKKCEASGFCYVNDIVLAILELLKYHQR 181

Qy    182 VLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLN 241
          |||:|||:|||||||||||||||||||||||||||:||||| |||||  |||||:|||||:
Db    182 VLYVDIDIHHGDGVEEAFYTTDRVMTVSFHKFGDYFPGTGDIRDVGHSKGKYYSLNVPLD 241

Qy    242 DGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLR 301
          ||:|||::|||:|:: ||||| | |||||||||||||||||||||:||||:|:|::|
Db    242 DGIDDESYQSLFKPIMGKVMEVFNPGAVVLQCGADSLSGDRLGCFNLSIKGHAECVKFMR 301

Qy    302 SYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP 361
          |:||||::||| |||||||||||||| ||:| |  :|:| |||:||||||||| ||
Db    302 SFNVPLLLLGGGGYTIRNVARCWCYETGVALGHELTDKMPPNEYYEYFGPDYTLHVAPSN 361

Qy    362 MENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSG 421
          ||| ||   :: |:: ||: || | |||||||  ||   : :: ||  :  |  |
Db    362 MENKNTRHQLDDIKSKLLDNLSKLRHAPSVQFQERPPEAELPEQDED--KENPDER---- 415

Qy    422 TATYESDSD---DDDKPL 436
              :::|||   :| |||
Db    416 ---HDADSDVEMNDAKPL 430


RESULT   4
US-08-528-255A-1
; Sequence 1, Application US/08528255A
; Patent No. 5659016
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  :  FURUKAWA, YOICHI
;    TITLE OF INVENTION:  RPDL PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
```

```
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/528,255A
;      FILING DATE:  September 14, 1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP6-227876
;      FILING DATE:  22-SEPTEMBER-1994
;      APPLICATION NUMBER:  JP7-183763
;      FILING DATE:  20-JULY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1335
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  482
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  human fetal lung cDNA library
US-08-528-255A-1


  Query Match             56.9%;  Score 1448.5;  DB 1;  Length 482;
  Best Local Similarity   58.3%;  Pred. No. 7.6e-152;
  Matches  261;  Conservative   79;  Mismatches   81;  Indels   27;  Gaps    4;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | :|:| |:|:  :|:||||||||||||||||| |:|:::| |:|::|| ||  |:| ::
Db       6 GTRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|: || |: |::| : |  : ::|||||||||||||||:||: | |||: :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVK 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  |||:|| |||||||||||||||||||||| |||||  :||||||||:||||||
Db     124 LNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||||||||:|::|||| :||:|| |||||:| || ||:|||: ::|:|
Db     184 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKP 243

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :: ||||::|| |||||||:|||||||||||||::|||| |: |::|:|:|:::||| ||
Db     244 VMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGY 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           |||||||| |||||:  |  |:||||:||||||||: ||: || | | || : :|:|:
Db     304 TIRNVARCRTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
             | |  | ||| || |  |  : : :|  |:                    | || ||
Db     364 QRLFENLRMLPHAPGVQMQAIPE-DAIPEESGDE-----------------DEDDPDKR 404

Qy     436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
           :      ::|:    |:|  |   |
Db     405 I------SICSSDKRIACEEEFSDSEEE 426
```

```
RESULT   5
US-08-717-365-1
; Sequence 1, Application US/08717365
; Patent No. 5763182
;  GENERAL INFORMATION:
;    APPLICANT:  NAKAMURA, YUSUKE
;    APPLICANT:  : FURUKAWA, YOICHI
;    TITLE OF INVENTION:  RPDL PROTEIN AND DNA
;    TITLE OF INVENTION:  ENCODING THE SAME
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLYNN, THIEL, BOUTELL & TANIS, P.C.
;      STREET:  2026 Rambling Road
;      CITY:  Kalamazoo
;      STATE:  Michigan
;      COUNTRY:  USA
;      ZIP:  49008-1699
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inches, 1.44 Mb storage
;      COMPUTER:  IBM PC/XT/AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.0
;      SOFTWARE:  WordPerfect 5.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,365
;      FILING DATE:  23-SEP-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/528 255
;      FILING DATE:  September 14, 1995
;      APPLICATION NUMBER:  JP6-227876
;      FILING DATE:  22-SEPTEMBER-1994
;      APPLICATION NUMBER:  JP7-183763
;      FILING DATE:  20-JULY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Terryence F. Chapman
;      REGISTRATION NUMBER:  32549
;      REFERENCE/DOCKET NUMBER:  Furuya Case 1335
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (616) 381-1156
;      TELEFAX:  (616) 381-5465
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  482
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    IMMEDIATE SOURCE:
;      LIBRARY:  human fetal lung cDNA library
US-08-717-365-1


  Query Match             56.9%;  Score 1448.5;  DB 1;  Length 482;
  Best Local Similarity   58.3%;  Pred. No. 7.6e-152;
  Matches  261;  Conservative   79;  Mismatches   81;  Indels   27;  Gaps    4;

Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
           | :|:| |:|:  :|:||||||||||||||||| |:|:::| |:|::|| ||  |:| ::
Db       6 GTRRKVCYYYDGDVGNYYYGQGHPMKPHRIRMTHNLLLNYGLYRKMEIYRPHKANAEEMT 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
           ::|| :|: || |: |::| : |  : ::|||||||||||||||:||: | |||: :|||
Db      66 KYHSDDYIKFLRSIRPDNMSEYS--KQMQRFNVGEDCPVFDGLFEFCQLSTGGSVASAVK 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
           ||:|  |||:|| |||||||||||||||||||||| |||||  :||||||||:||||||
Db     124 LNKQQTDIAVNWAGGLHHAKKSEASGFCYVNDIVLAILELLKYHQRVLYIDIDIHHGDGV 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
           |||||||||||||||||:|::|||| :||:|| |||||:| || ||:|||: ::|:|
Db     184 EEAFYTTDRVMTVSFHKYGEYFPGTGDLRDIGAGKGKYYAVNYPLRDGIDDESYEAIFKP 243
```

```
Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
           :: ||||::|| ||||||||:|||||||||||||::|||| |: |:::|:|:|:::||| ||
Db     244 VMSKVMEMFQPSAVVLQCGSDSLSGDRLGCFNLTIKGHAKCVEFVKSFNLPMLMLGGGGY 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
           |||||||| |||||: | |:||||:||||||||: ||: || | | || : :|:|:
Db     304 TIRNVARCRTYETAVALDTEIPNELPYNDYFEYFGPDFKLHISPSNMTNQNTNEYLEKIK 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
             | | |  | ||| || |  |  : : :|  |:                 | || ||
Db     364 QRLFENLRMLPHAPGVQMQAIPE-DAIPEESGDE----------------DEDDPDKR 404

Qy     436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
           :        ::|:    |:|  |   |
Db     405 I------SICSSDKRIACEEEFSDSEEE 426


RESULT   6
US-09-282-305-4
; Sequence 4, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 351
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-4


  Query Match             50.3%;  Score 1280.5;  DB 4;  Length 351;
  Best Local Similarity   53.5%;  Pred. No. 1.9e-133;
  Matches  252;  Conservative   39;  Mismatches   55;  Indels  125;  Gaps    6;

Qy       1 MEADESGISLPS---GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHL 57
           | |   | ||||   | |  :|||||||||:|||||||||||||
Db       1 MAASGEGASLPSPAGGEDAHRRRVSYFYEPSIGDYYYGQGHPMKP--------------- 45

Qy      58 HRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDG 117

Db      46 ------------------------------------------------------------ 45

Qy     118 LFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLK 177
                                          ||||||||||||||||||| ||||||
Db      46 ----------------------------------HHAKKSEASGFCYVNDIVLAILELLK 71

Qy     178 MFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALN 237
             :||||:|||||||||||||:||:|||||||||:|||||||||| |||| :||:||||
Db      72 FHRRVLYVDIDVHHGDGVEEAFFTTNRVMTVSFHKYGDFFPGTGHITDVGAAEGKHYALN 131

Qy     238 VPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCL 297
           |||:||:|| :|| ||: :|:|||||||||: ||||||||||:||||||||||||||||||
Db     132 VPLSDGIDDTTFRGLFQCIIKKVMEVYQPDVVVLQCGADSLAGDRLGCFNLSVKGHADCL 191

Qy     298 RFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHV 357
           |||||||||:||||| ||||||||||||||||||||||||||||:|:||||||||||:
Db     192 RFLRSYNVPMMVLGGGGYTIRNVARCWCYETAVAVGVEPDNKLPYNDYYEYFGPDYTLHI 251

Qy     358 DPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRP--K 415
             |  :||||| ||:| |:| :|| || : | || ||   |      :| |:||: ||  :
Db     252 QPKSVENLNTTKDLENIKNMILENLSKIEHVPSTQFHDRPSDPEAPEEKEEDMDKRPPQR 311

Qy     416 PRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDV 466
             | ||| | |:||::| |              :  :| |   :| |: |: |: |:
Db     312 SRLWSGGA-YDSDTEDPDSL---------KSEGKDVTANFQMKDE-PKDDL 351


RESULT   7
US-09-282-305-10
; Sequence 10, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
;  APPLICANT:  Baldwin, Donald A.
;  APPLICANT:  Briggs, Steven P.
;  APPLICANT:  Crane, Virginia C.
;  TITLE OF INVENTION:  Maize Histone Deacetylases And Their Uses
;  FILE REFERENCE:  5718-44,
;  CURRENT APPLICATION NUMBER: US/09/282,305
;  CURRENT FILING DATE:  1999-03-31
;  PRIOR APPLICATION NUMBER:  60/080,563
;  PRIOR FILING DATE:  1998-04-03
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 432
;   TYPE: PRT
;   ORGANISM: Zea mays
US-09-282-305-10


  Query Match             49.1%;  Score 1249.5;  DB 4;  Length 432;
  Best Local Similarity   55.6%;  Pred. No. 7.5e-130;
  Matches  235;  Conservative   64;  Mismatches  107;  Indels   17;  Gaps    5;

Qy      23 YFYEPTIGD---YYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHS 79
           ||  | :||    |:|  |||||||: | | |:: | ||:::|| ||  |   :: :|||
Db       8 YFITPIVGDVGNVYFGPNHPMKPHRLCMTHHLVLSYGLHQKMEIYRPHKAYPIELAQFHS 67

Qy      80 PEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQ 139
             :||:||  ::|:|      |  | |:|:|||||||| ||:||:  |||:: || :|| :
Db      68 ADYVEFLHRITPDS--QHLYASELTRYNLGEDCPVFDNLFEFCQIYAGGTLDAARRLNHK 125

Qy     140 DADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAF 199
             |||||| |||||||| |||||||:||:|||||||||   ||||||||||||||||||
Db     126 ICDIAINWAGGLHHAKKCEASGFCYINDLVLGILELLKYHARVLYIDIDVHHGDGVEEAF 185

Qy     200 YTTDRVMTVSFHKFGD-FFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQ 258
           | ||||||||||:|| |||||| |:|:|  :|||||:|:|| ||:|| ||   |: :|
Db     186 YFTDRVMTVSFHKYGDLFFPGTGDIKDIGEREGKYYAINIPLKDGIDDTSFTRPFKTIIA 245

Qy     259 KVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTIR 318
           ||:| | | |:|||||||||: ||||||||||::|||:|::|:: :|:|||:| || |||
Db     246 KVVETYLPGAIVLQCGADSLARDRLGCFNLSIEGHAECVKFVKKFNIPLLVTGGGGYTKE 305

Qy     319 NVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNTL 378
           |||||| || | :  |  |::| ||| ||| ||||| |    |:|||:   :  |:  :
Db     306 NVARCWAVETGVLLDTELPNEIPKNEYIEYFAPDYTLKVPNLNMDNLNSKTYLSSIKVQV 365

Qy     379 LEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTAT---------YESDS 429
           :| |  : ||| || |  ||   : |  ||::  |  |     |           || |:
Db     366 MESLRYIQHAPGVQMQEVPPDFYIPDFDEDELD--PDERVDQHTQDKQIHRDDEYYEGDN 423

Qy     430 DDD 432
           |:|
Db     424 DND 426


RESULT   8
US-09-446-504-80
; Sequence 80, Application US/09446504
; Patent No. 6218150
; GENERAL INFORMATION:
;  APPLICANT:  UEMORI, Takashi
;  APPLICANT:  SATO, Yoshimi
;  APPLICANT:  FUJITA, Tomoko
```

```
;  APPLICANT:  MIYAKE, Kazue
;  APPLICANT:  MUKAI, Hiroyuki
;  APPLICANT:  ASADA, Kiyozo
;  APPLICANT:  KATO, Ikunoshin
;  TITLE OF INVENTION:  DNA POLYMERASE-RELATED FACTORS
;  FILE REFERENCE:  1422-408PCT
;  CURRENT APPLICATION NUMBER: US/09/446,504
;  CURRENT FILING DATE:  1999-12-23
;  PRIOR APPLICATION NUMBER:  PCT/JP98/02845
;  PRIOR FILING DATE:  1998-06-24
;  PRIOR APPLICATION NUMBER:  JP 9-187496
;  PRIOR FILING DATE:  1997-06-26
;  PRIOR APPLICATION NUMBER:  JP 9-320692
;  PRIOR FILING DATE:  1997-11-27
;  NUMBER OF SEQ ID NOS:  92
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 80
;   LENGTH: 335
;   TYPE: PRT
;   ORGANISM: Pyrococcus furiosus
US-09-446-504-80


  Query Match             9.7%;  Score 247.5;  DB 4;  Length 335;
  Best Local Similarity   22.1%;  Pred. No. 7.3e-19;
  Matches   68;  Conservative   64;  Mismatches  139;  Indels   37;  Gaps    8;

Qy      22 SYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPE 81
           | ||       :   : ||  | |: :  | :     |  |:|   |:  : : : : :  |  :
Db       4 SIFYSKKFNLHRPSEYHPENPKRLEIVLSKVRELGLEERIE--EPNPVEETFVEKIHDRD 61

Qy      82 YVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDA 141
           |::|:           :  : : | :   |  |  | :     : | :  ||:
Db      62 YINFVKE---------AVEKGITRLD--PDTYVSPGTWSAALLALGAARSAALS------ 104

Qy     142 DIAINWGG---------GLHHAKKSEAS-----GFCYVNDIVLGILELLKM-FKRVLYID 186
             |:::||         | |  ::  |       |||  |:    :: | :    :|: ||
Db     105 --ALHYGGLHMALVRPPGHHAGRRGRAMGAPTLGFCIFNNAASAVVTLKEEGVGKVVVID 162

Qy     187 IDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDD 246
           | |||:| :| |:    |: :  |:  | :||:| : :||        :|:|:     |
Db     163 FDAHHGNGTQEIFWNDPDVIHIDLHE-RDIYPGSGDVSEVGGSNAYGSKINLPMPHYSGD 221

Qy     247 ESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVP 306
             :  ::  ::  ::|  :|: :|:  | |   || |    |:   ::      |  | :
Db     222 GDYIYVWDEIVLPIVEEVKPKVIVISAGFDGFKGDGLTTLRLTESFYSYAGATLNKYPLA 281

Qy     307 LMVLGGEG 314
            :: || |
Db     282 FILEGGYG 289


RESULT   9
US-09-144-085-2
; Sequence 2, Application US/09144085
; Patent No. 6280999
; GENERAL INFORMATION:
;  APPLICANT: Gustafsson, Claes
;  APPLICANT: Betlach, Mary C.
;  APPLICANT: Ashley, Gary
;  APPLICANT: Julien, Bryan
;  APPLICANT: Ziermann, Rainer
;  TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
;  TITLE OF INVENTION: THEREFOR
;  FILE REFERENCE: 30062-20020.20
;  CURRENT APPLICATION NUMBER: US/09/144,085
;  CURRENT FILING DATE: 1998-08-31
;  EARLIER APPLICATION NUMBER: 09/010,809
;  EARLIER FILING DATE: 1998-01-22
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 6095
;   TYPE: PRT
;   ORGANISM: Sorangium cellulosum
US-09-144-085-2


  Query Match             4.1%;  Score 104.5;  DB 4;  Length 6095;
  Best Local Similarity   20.1%;  Pred. No. 0.68;
  Matches  109;  Conservative   59;  Mismatches  193;  Indels  181;  Gaps   24;

Qy       2 EADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
           | | :|  | :|  |                     |  :  | :| ||  :  |  |  |
Db    1426 ELDPAGTVLVTGGTGE-----------------LGQAIAAHLVR-AHG-VRHLVLTSRR 1465

Qy      62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMG----DPSAARNLRRFNVGED------ 111
             :  |          : | |  |  :  |::     | |    : |   | |
Db    1466 GLEAPG----------ARELVQSLEELGAETVTVAACDVSKREEVARVLAGIDAARPLS 1514

Qy     112 --------------------------CPVFDG---LFDFCR----------ASAGGSIGA 132
                                       |  ||   | :  |          :|| |: ||
Db    1515 AVLHLAGVLDDGVLTAQTAERLSRVLAPKVDGALHLHELTRELDLAAFVLFSSAAGTFGA 1574

Qy     133 AVKLNRQ------DADIAINWGGGLHHAKKSEASGFCYVNDI----VLGILELLKMFKRV 182
           | : |         ||  |   ||||  |  | | ||      :     ||  || :| :|
Db    1575 AGQSNYAAANTFLDALAAHRRGGGL--AATSLAWGFWTQAGVGMTAHLGEAELSRM-RRN 1631

Qy     183 LYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLND 242
           :: : |  |  : :|  :       |  |                 |:           |
Db    1632 GFVPMPVEEGLALLDAALSRPEASLVPVH-----------LDL----------AQLQR 1668

Qy     243 GMDDES-----FRSLFRPLIQKVMEVYQPEAVVLQCGADSL-SGDRLGCFNLSVKGHADC 296
           |::         ||:| || ::|     : :|  |:    :|   :||       |:|
Db    1669 GLESSGELPALFRALLRPSLRKASSATRRDASALRERLSALPEAERLNALVELVRGEVAA 1728

Qy     297 LRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGV------EPDNKLPYNEYFEYFG 350
           :  |:          ||      | :    :: :|| :        :  ||    |:|
Db    1729 VAGLQR---------GEAVAADQVLKELGLDSLMAVALRNRLTSRTETSLPATLVFDY-- 1777

Qy     351 PDYTLHVDPSP-----------MENLNTPKDMERIRN-----------TLLEQLSGLIHA 388
                 |:|             |   :   |:|            ::  :| | :
Db    1778 --------PTPRAIAELLLKQAFSGLQVKEARARVRRRAGKDEPIAIVSMACRLPGGVAT 1829

Qy     389 PSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKPLHGYSCRGGATTD 448
           |   :       |:| | :| :|    |  | |   |: | :   |    |:  ||   |
Db    1830 PDDYW-------RLLAEGKDAIE--GLPARWDGFEVYDPDPEAAGK---SYAREGGFVRD 1877

Qy     449 RD 450
           ||
Db    1878 ID 1879


RESULT  10
US-09-428-517-2
; Sequence 2, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
;  APPLICANT: Betlach, Mary C.
;  APPLICANT: Shah, Sanjay Krishnakant
;  APPLICANT: McDaniel, Robert
;  APPLICANT: Tang, Li
;  TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
;  FILE REFERENCE: 30062-20029.00
;  CURRENT APPLICATION NUMBER: US/09/428,517
;  CURRENT FILING DATE: 1999-10-28
;  EARLIER APPLICATION NUMBER: 60/120,254
;  EARLIER FILING DATE: 1999-02-16
;  EARLIER APPLICATION NUMBER: 60/106,100
;  EARLIER FILING DATE: 1998-10-29
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4150
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
```

```
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Recombinant
;   OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2


  Query Match             4.0%;  Score 101.5;  DB 4;  Length 4150;
  Best Local Similarity   24.0%;  Pred. No. 0.77;
  Matches   76;  Conservative   33;  Mismatches  111;  Indels   97;  Gaps   17;

Qy     189 VHHGDGVEE-------AFYTT--DRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVP 239
           |  |||:          :|     | |:|          |  ||  : ||:   : || :|
Db     823 VRFGDGIRALGKLGTGSFLEVGPDGVLTAMARACVTAAPEPGHRGEQGADADAHTALLLP 882

Qy     240 -LNDGMDDESFRSLFRPLIQKVME---------------------VYQPEAVVLQCGA- 275
           |  | |:   |||   : :  :                       :| |:  |  |
Db     883 ALRRGRDEA--RSLTEAVARLHLHGVPMDWTSVLGGDVSRVPLPTYAFQRESHWLPSGEA 940

Qy     276 ------DSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETA 329
                |: ||   |     |      | |  :||:     |||                :
Db     941 HPRPADDTESG--TGRTEASPPRPHDVLHLVRSH--AAAVLG---------------HS 980

Qy     330 VAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAP 389
           |  ::||         |   |             ::|   :  :|:   |     ||
Db     981 RAERIDPDRA------FRDLG----------FDSLTALELRDRLDTAL-----GLRLPS 1018

Qy     390 SVQFQHTPP--VNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKPLH--GYSCR--G 443
           || | |  |  : | |   : |   ||:|   :||   |   | ||:|:   | :||  |
Db    1019 SVLFDHPSPGALARFL---QGDDTRRPEPGKTNGTRATEPGPDPDDEPIAIVGMACRFPG 1075

Qy     444 GATTDRD-----STGED 455
           | |:  |     : |||
Db    1076 GVTSPEDLWRLLAAGED 1092


RESULT  11
US-08-296-362-2
; Sequence 2, Application US/08296362
; Patent No. 5691306
;  GENERAL INFORMATION:
;    APPLICANT:  Bergeron, John J.M.
;    APPLICANT:  Thomas, David Y.
;    APPLICANT:  Wada, Ikuo
;    TITLE OF INVENTION:  METHODS OF DETECTION AND TREATMENT OF
;    TITLE OF INVENTION:  PROTEIN TRAFFICKING DISORDERS AND INCREASING SECRETORY
;    TITLE OF INVENTION:  PROTEIN PRODUCTION
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SEED and BERRY
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/296,362
;      FILING DATE:  25-AUG-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Deehr, Manya S.
;      REGISTRATION NUMBER:  37,120
;      REFERENCE/DOCKET NUMBER:  690066.401C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;      TELEX:  3723836
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  593 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-296-362-2


  Query Match             3.7%;  Score 93;  DB 1;  Length 593;
  Best Local Similarity   21.4%;  Pred. No. 0.26;
  Matches   70;  Conservative   33;  Mismatches   96;  Indels  128;  Gaps   18;

Qy     213 FGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQ 272
           | | |       | |   |  : |:    |  |||             :  |  :  | :
Db      72 FADSF-------DRGTLSG--WILSKAKKDDTDDE------------IAKYDGKWEVDE 109

Qy     273 CGADSLSGDRLGCFNLSVKGHADCLRFLRSY---NVPLMVLGGEGYTIRNVARCWCYETA 329
             | ||:          | ||   :  : :     ||:|                ||
Db     110 MKETKLPGDKGLVLMSRAKHHAISAKLNKPFLFDTKPLIVQ--------------YEVN 154

Qy     330 VAVGVE-----------PD-------NKLPYNEYF--EYFGPDYTLHV---DPSPMENL 365
             |:|             |:       :| ||   |  :  | || ||      :|   :
Db     155 FQNGIECGGAYVKLLSKTPELNLDQFHDKTPYTIMFGPDKCGEDYKLHFIFRHKNPKTGV 214

Qy     366 NTPKDMER-----------------IRN-----TLLEQLSGLIHAPSVQFQHTPPVN 400
             |  :|                 | |       |::|   :::: ::    |||||
Db     215 YEEKHAKRPDADLKTYFTDKKTHLYTLILNPDNSFEILVDQ--SIVNSGNLLNDMTPPVN 272

Qy     401 --RVLDEPED----DMETRP--------KPRXWSGTATYESDSDDDDKPLHGYSCRGGAT 446
             | :::|||     | : ||        ||  |           ::| |
Db     273 PSREIEDPEDQKPEDWDERPKIPDPDAVKPDDW----------NEDAP--------AKI 313

Qy     447 TDRDSTGEDEMDDDNPE----PDVNPP 469
           | ::|  |   || ||    ||   |
Db     314 PDEEATKPDGWLDDEPEYVPDPDAEKP 340


RESULT  12
US-08-867-611-2
; Sequence 2, Application US/08867611
; Patent No. 6172189
;  GENERAL INFORMATION:
;    APPLICANT:  DEVARE, SUSHIL G
;    APPLICANT:  DESAI, SURESH M
;    APPLICANT:  CASEY, JAMES M
;    APPLICANT:  DAILEY, STEPHEN H
;    APPLICANT:  DAWSON, GEORGE J
;    APPLICANT:  GUTIERREZ, ROBIN A
;    APPLICANT:  LESNIEWSKI, RICHARD R
;    APPLICANT:  STEWART, JAMES L
;    APPLICANT:  RUPPRECHT, KEVIN R
;    TITLE OF INVENTION:  HEPATITIS C ASSAY UTILIZING RECOMBINANT
;    TITLE OF INVENTION:  ANTIGENS
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ABBOTT LABORATORIES
;      STREET:  ONE ABBOTT PARK ROAD, CHAD377/AP6D2
;      CITY:  ABBOTT PARK
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/867,611
;      FILING DATE:  02-JUN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/646,757
;      FILING DATE:
;      APPLICATION NUMBER:  US/08/179,896
```

```
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/572,822
;      FILING DATE:  24-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/614,069
;      FILING DATE:  07-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/748,561
;      FILING DATE:  21-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/748,565
;      FILING DATE:  21-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/748,566
;      FILING DATE:  21-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  POREMBSKI, PRISCILLA E
;      REGISTRATION NUMBER:  33,207
;      REFERENCE/DOCKET NUMBER:  4834.US.P6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708-937-6365
;      TELEFAX:  708-937-9556
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  396 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-867-611-2


  Query Match             3.6%;  Score 91;  DB 4;  Length 396;
  Best Local Similarity   22.4%;  Pred. No. 0.22;
  Matches   59;  Conservative   33;  Mismatches   83;  Indels   88;  Gaps   14;

Qy     194 GVEEAFYTTDR--------------VMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVP 239
           | |     ||                 || : |:      ||  : :|  ||  :    |
Db      41 GAERIIVATDHEDVARAVEAAGGEVCMTRADHQ-----SGTERLAEV-VEKCAFSDDTVI 94

Qy     240 LNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRF 299
           :|   |:        |:|        |  :: |  ||:|:  ::|    |:|  |     |
Db      95 VNVQGDE--------PMI--------PATIIRQV-ADNLAQRQVGMATLAVPIHNAEEAF 137

Qy     300 LRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHV-- 357
             : |   :||  ||| :        | :   :  : |    : |  |  | :: |:
Db     138 --NPNAVKVVLDAEGYAL--------YFSRATIPWDRDR---FAEGLETVGDNFLRHLGI 184

Qy     358 --------------DPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQF---QHTPPVN 400
                         |||:|::            :|||  | :   :     |  |
Db     185 YGYRAGFIRRYVNWQPSPLEHIE----------MLEQLRVLWYGEKIHVAVAQEVPGTG 233

Qy     401 RVLDEPED------DMETRPKPR 417
             :| |||       | | |||:
Db     234 --VDTPEDLDPSTNSMSTNPKPQ 254


RESULT  13
PCT-US92-06965A-7
; Sequence 7, Application PC/TUS9206965A
;  GENERAL INFORMATION:
;    APPLICANT:  DEVARE, S.
;    APPLICANT:  DESAI, S.
;    APPLICANT:  DAILEY, S.
;    TITLE OF INVENTION:  HCV SYNTHETIC PEPTIDE FROM NS1 REGION
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ABBOTT LABORATORIES
;      STREET:  ONE ABBOTT PARK ROAD
;      CITY:  ABBOTT PARK
;      STATE:  ILLINOIS
;      COUNTRY:  U.S.
;      ZIP:  60065-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/06965A
;      FILING DATE:  19920821
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  POREMBSKI, PRISCILLA E.
;      REGISTRATION NUMBER:  33,207
;      REFERENCE/DOCKET NUMBER:  4834PC.02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708-937-6365
;      TELEFAX:  708-937-9556
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  396 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US92-06965A-7


  Query Match             3.6%;  Score 91;  DB 5;  Length 396;
  Best Local Similarity   22.4%;  Pred. No. 0.22;
  Matches   59;  Conservative   33;  Mismatches   83;  Indels   88;  Gaps   14;

Qy     194 GVEEAFYTTDR--------------VMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVP 239
           | |     ||                 || : |:      ||  : :|  ||  :    |
Db      41 GAERIIVATDHEDVARAVEAAGGEVCMTRADHQ-----SGTERLAEV-VEKCAFSDDTVI 94

Qy     240 LNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRF 299
           :|   |:        |:|        |  :: |  ||:|:  ::|    |:|  |     |
Db      95 VNVQGDE--------PMI--------PATIIRQV-ADNLAQRQVGMATLAVPIHNAEEAF 137

Qy     300 LRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHV-- 357
             : |   :||  ||| :        | :   :  : |    : |  |  | :: |:
Db     138 --NPNAVKVVLDAEGYAL--------YFSRATIPWDRDR---FAEGLETVGDNFLRHLGI 184

Qy     358 --------------DPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQF---QHTPPVN 400
                         |||:|::            :|||  | :   :     |  |
Db     185 YGYRAGFIRRYVNWQPSPLEHIE----------MLEQLRVLWYGEKIHVAVAQEVPGTG 233

Qy     401 RVLDEPED------DMETRPKPR 417
             :| |||       | | |||:
Db     234 --VDTPEDLDPSTNSMSTNPKPQ 254


RESULT  14
US-08-291-896-2
; Sequence 2, Application US/08291896
; Patent No. 5480800
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA fragment carrying the gene encoding the
;    TITLE OF INVENTION:  enzyme for fragmenting N-acetylheparosan and the adjacent
;    TITLE OF INVENTION:  sequences permitting its expression, recombinant enzyme
;    TITLE OF INVENTION:  intended for fragmenting N-acetylheparosan and its use.
;    NUMBER OF SEQUENCES:  7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/291,896
;      FILING DATE:  17-AUG-1994
;      CLASSIFICATION:  435
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  820 amino acids
```

```
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-291-896-2


  Query Match             3.5%;  Score 88;  DB 1;  Length 820;
  Best Local Similarity   19.8%;  Pred. No. 1.6;
  Matches   69;  Conservative   36;  Mismatches  123;  Indels  120;  Gaps   17;

Qy     164 YVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHI 223
           ||:  : |::     |         : :| |   | | |          | |       |
Db     163 YVDAAIAGVIGGFGWF---------IQYGSG---AVYRT----------FQDKMRDGVSI 200

Qy     224 RDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAV-------VLQCGAD 276
           :| ||: |        |||  |       |   :     |: || |       : :||
Db     201 KDFGAQNG-------ILNDNKD------AFTKSLHSFSSVFVPEGVFNTSLVSLSRCGLY 247

Qy     277 SLSG------DRLG---CFNL-----------------SVKGHADCLRFLRSYNVPLMV 309
             |      || |    ||:                 ||:||             :
Db     248 GTGGGTIKQYDRDGNHLVFNMPDGGMLSTLTIMGNKSDDSVQGH------------QVSF 295

Qy     310 LGGEGYTIRNVARCWCYETAVAVGVEPDNKLP--------YNEYFEY----------FGP 351
           ||   :::|:          ::   ||| :|          ||  :          |
Db     296 SGGHDVSVKNIRFTNTRGPGFSLIAYPDNGIPSGYIVRDIRGEYLGFANNKKAGCVLFDS 355

Qy     352 DYTLHVDPSPMEN--------LNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVL 403
             :|     |         | |      : | : |:   ::: :    | | | ::
Db     356 SQNTLIDGVIARNYPQFGAVELKTAAKYNIVSNVIGEECQHVVYNGTE--TETAPTNNII 413

Qy     404 DEPEDDMETRPKPRX---WSGTATYESD-----SDDDDKPLHGYSCRG 443
             |   ||         ||     ||      |: | |  || : :|
Db     414 SSV---MANNPKYAAVVVGKGTGNLISDVLVDYSESDAKQAHGVTVQG 458


RESULT  15
US-08-485-278-2
; Sequence 2, Application US/08485278
; Patent No. 5820857
;  GENERAL INFORMATION:
;    APPLICANT:  Legoux, Richard
;    APPLICANT:  Lelong, Philippe
;    APPLICANT:  Salome, Marc Louis Victor
;    TITLE OF INVENTION:  DNA fragment carrying the gene encoding the enzyme for
;    TITLE OF INVENTION:  fragmenting N-acetylheparosan and the adjacent sequences permi
;    TITLE OF INVENTION:  recombinant enzyme intended for fragmenting N-acetylheparosan
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jacobson, Price, Holman & Stern
;      STREET:  400 Seventh St. N.W.
;      CITY:  Washington D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,278
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 93/10050
;      FILING DATE:  17-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/291,896
;      FILING DATE:  17-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Player, William E.
;      REGISTRATION NUMBER:  31,409
;      REFERENCE/DOCKET NUMBER:  P58114NA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 638-6666
;      TELEFAX:  (202) 393-5350
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  820 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-485-278-2


  Query Match             3.5%;  Score 88;  DB 2;  Length 820;
  Best Local Similarity   19.8%;  Pred. No. 1.6;
  Matches   69;  Conservative   36;  Mismatches  123;  Indels  120;  Gaps   17;

Qy     164 YVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHI 223
           ||:  : |::     |         : :| |   | | |          | |       |
Db     163 YVDAAIAGVIGGFGWF---------IQYGSG---AVYRT----------FQDKMRDGVSI 200

Qy     224 RDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAV-------VLQCGAD 276
           :| ||: |        |||  |       |   :     |: || |       : :||
Db     201 KDFGAQNG-------ILNDNKD------AFTKSLHSFSSVFVPEGVFNTSLVSLSRCGLY 247

Qy     277 SLSG------DRLG---CFNL-----------------SVKGHADCLRFLRSYNVPLMV 309
             |      || |    ||:                 ||:||             :
Db     248 GTGGGTIKQYDRDGNHLVFNMPDGGMLSTLTIMGNKSDDSVQGH------------QVSF 295

Qy     310 LGGEGYTIRNVARCWCYETAVAVGVEPDNKLP--------YNEYFEY----------FGP 351
           ||   :::|:          ::   ||| :|          ||  :          |
Db     296 SGGHDVSVKNIRFTNTRGPGFSLIAYPDNGIPSGYIVRDIRGEYLGFANNKKAGCVLFDS 355

Qy     352 DYTLHVDPSPMEN--------LNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVL 403
             :|     |         | |      : | : |:   ::: :    | | | ::
Db     356 SQNTLIDGVIARNYPQFGAVELKTAAKYNIVSNVIGEECQHVVYNGTE--TETAPTNNII 413

Qy     404 DEPEDDMETRPKPRX---WSGTATYESD-----SDDDDKPLHGYSCRG 443
             |   ||         ||     ||      |: | |  || : :|
Db     414 SSV---MANNPKYAAVVVGKGTGNLISDVLVDYSESDAKQAHGVTVQG 458


Search completed: April 26, 2002, 17:51:27
Job time: 35 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 26, 2002, 17:52:13 ; Search time 43 Seconds
(without alignments)
811.361 Million cell updates/sec

Title: US-09-645-337-4
Perfect score: 2544
Sequence: 1 MEADESGISLPSGPDGRKRR.....TGEDEMDDDNPEPDVNPPSS 471

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2542	99.9	471	22	AAB67812	Amino acid sequenc
2	2535	99.6	471	21	AAG05187	Arabidopsis thalia
3	2535	99.6	471	21	AAG39372	Arabidopsis thalia
4	2318	91.1	432	21	AAG05188	Arabidopsis thalia
5	2318	91.1	432	21	AAG39373	Arabidopsis thalia
6	2279	89.6	425	21	AAG05189	Arabidopsis thalia
7	2279	89.6	425	21	AAG39374	Arabidopsis thalia
8	1734	68.2	458	20	AAY28797	Maize histone deac
9	1561.5	61.4	501	22	AAB67811	Amino acid sequenc
10	1561.5	61.4	501	22	AAB67165	Arabidopsis histon
11	1558.5	61.3	473	21	AAY58829	Soybean histone de
12	1551.5	61.0	517	20	AAY28800	Maize histone deac
13	1520	59.7	439	20	AAY28799	Maize histone deac
14	1498	58.9	493	21	AAY58828	Rice histone deace
15	1462.5	57.5	482	18	AAW29324	A novel histone de
16	1462.5	57.5	482	22	AAB67164	Human histone deac
17	1462.5	57.5	482	22	AAB49954	Human histone deac
18	1462.5	57.5	488	21	AAB56985	Human prostate can
19	1448.5	56.9	482	17	AAR88919	RPDL, a human foet
20	1445.5	56.8	488	22	AAB49955	Human histone deac
21	1348.5	53.0	370	22	AAU00243	Human histone deac
22	1328.5	52.2	433	22	AAB67166	Yeast histone deac
23	1292.5	50.8	428	22	AAB49956	Human histone deac
24	1280.5	50.3	351	20	AAY28798	Maize histone deac
25	1249.5	49.1	432	20	AAY28801	Maize histone deac
26	813.5	32.0	415	22	AAM40374	Human polypeptide
27	813	32.0	377	21	AAY83027	Human histone deac
28	813	32.0	377	22	AAB84361	Amino acid sequenc
29	594	23.3	180	22	AAB70496	Saccharomyces cere
30	561.5	22.1	296	22	AAM42160	Human polypeptide
31	465.5	18.3	375	22	AAU00241	Cys75Ser/Cys77Ser
32	460	18.1	125	18	AAW29330	A partial deacetyl
33	459	18.0	133	18	AAW29329	A partial deacetyl
34	457.5	18.0	375	22	AAU00240	Wild-type histone
35	453.5	17.8	375	22	AAU00242	Tyr297Phe mutant o
36	328.5	12.9	120	21	AAY58830	Wheat histone deac
37	309.5	12.2	74	22	AAB48580	Human ovarian canc
38	309	12.1	69	18	AAW29326	HDx protein family
39	309	12.1	398	21	AAG07246	Arabidopsis thalia
40	302	11.9	364	21	AAG07248	Arabidopsis thalia
41	302	11.9	377	21	AAG07247	Arabidopsis thalia
42	295	11.6	80	18	AAW29331	A partial deacetyl
43	289	11.4	468	21	AAG51592	Arabidopsis thalia
44	289	11.4	499	21	AAG51591	Arabidopsis thalia
45	289	11.4	552	21	AAG51590	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAB67812
ID AAB67812 standard; Protein; 471 AA.
XX
AC AAB67812;
XX
DT 29-JUN-2001 (first entry)
XX
DE Amino acid sequence of a histone deacetylase designated AtRPD3B.
XX
KW Histone deacetylase; AtRPD3B; RPD3; gene expression; transgenic plant;
KW HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS Arabidopsis thaliana.
XX
PN CA2316036-A1.
XX
PD 27-FEB-2001.
XX
PF 24-AUG-2000; 2000CA-2316036.
XX
PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX
DR WPI; 2001-258457/27.
DR N-PSDB; AAF80351.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT -

XX
PS Claim 10; Fig 1B; 91pp; English.
XX
CC The present sequence represents Arabidopsis thaliana histone deacetylase
CC designated AtRPD3B. The protein is homologous to yeast RPD3 and HDA1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operative
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.
XX
SQ Sequence 471 AA;

Query Match 99.9%; Score 2542; DB 22; Length 471;
Best Local Similarity 99.8%; Pred. No. 6.2e-263;
Matches 470; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 meadesgislpsgpdgrkrrvsyfyeptigdyyygqghpmkphrirmahsliihyhlhrr 60

Qy     61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 leisrpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrfnvgedcpvfdglfd 120

Qy    121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 fcrasaggsigaavklnrqdadiainwgggLhhakkseasgfcyvndivlgilellkmfk 180

Qy    181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 rvlyididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdvgaekgkyyalnvpl 240

Qy    241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ndgmddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcfnlsvkghadclrfl 300

Qy    301 RSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 rsynvplmvlggegytirnvarcwcyetavavgvepdnklpyneyfeyfgpdytlhvdps 360

Qy    361 PMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWS 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    361 pmenlntpkdmerirntlleqlsglihapsvqfqhtppvnrvldepeddmetrpkpriws 420

Qy    421 GTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gtatyesdsddddkplhgyscrggattdrdstgedemdddnpepdvnppss 471
```

RESULT 2
AAG05187
ID AAG05187 standard; Protein; 471 AA.
XX
AC AAG05187;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1499.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.

XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.6%; Score 2535; DB 21; Length 471;
Best Local Similarity 99.6%; Pred. No. 3.5e-262;
Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 meadesgislpsgpdgrkrrvsyfyeptigdyyygqghpmkphrirmahsliihyhlhrr 60

Qy      61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 leisrpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrfnvgedcpvfdglfd 120

Qy     121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 fcrasaggsigaavklnrqdadiainwggglhhakkseasgfcyvndivlgilellkmfk 180

Qy     181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 rvlyididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdvgaekgkyyalnvpl 240

Qy     241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ndgmddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcfnlsvkghadclrfl 300

Qy     301 RSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     301 rsynvplmvlggggytirnvarcwcyetavavgvepdnklpyneyfeyfgpdytlhvdps 360

Qy     361 PMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWS 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     361 pmenlntpkdmerirntlleqlsglihapsvqfqhtppvnrvldepeddmetrpkpriws 420
```

```
Qy    421 GTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gtatyesdsdddddkplhgyscrggattdrdstgedemdddnpepdvnppss 471


RESULT   3
AAG39372
ID   AAG39372 standard; Protein; 471 AA.
XX
AC   AAG39372;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 48705.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
PR   16-JUL-1999;   99US-0144085.
PR   16-JUL-1999;   99US-0144086.
PR   19-JUL-1999;   99US-0144325.
PR   19-JUL-1999;   99US-0144331.
PR   19-JUL-1999;   99US-0144332.
PR   19-JUL-1999;   99US-0144333.
PR   19-JUL-1999;   99US-0144334.
PR   19-JUL-1999;   99US-0144335.
PR   20-JUL-1999;   99US-0144352.
PR   20-JUL-1999;   99US-0144632.
PR   20-JUL-1999;   99US-0144884.
PR   21-JUL-1999;   99US-0144814.
PR   21-JUL-1999;   99US-0145086.
PR   21-JUL-1999;   99US-0145088.
PR   22-JUL-1999;   99US-0145085.
PR   22-JUL-1999;   99US-0145087.
PR   22-JUL-1999;   99US-0145089.
PR   22-JUL-1999;   99US-0145192.
PR   23-JUL-1999;   99US-0145145.
PR   23-JUL-1999;   99US-0145218.
PR   23-JUL-1999;   99US-0145224.
PR   26-JUL-1999;   99US-0145276.
PR   27-JUL-1999;   99US-0145913.
PR   27-JUL-1999;   99US-0145918.
PR   27-JUL-1999;   99US-0145919.
PR   28-JUL-1999;   99US-0145951.
PR   02-AUG-1999;   99US-0146386.
PR   02-AUG-1999;   99US-0146388.
PR   02-AUG-1999;   99US-0146389.
PR   03-AUG-1999;   99US-0147038.
PR   04-AUG-1999;   99US-0147204.
PR   04-AUG-1999;   99US-0147302.
PR   05-AUG-1999;   99US-0147192.
PR   05-AUG-1999;   99US-0147260.
PR   06-AUG-1999;   99US-0147303.
PR   06-AUG-1999;   99US-0147416.
PR   09-AUG-1999;   99US-0147493.
PR   09-AUG-1999;   99US-0147935.
PR   10-AUG-1999;   99US-0148171.
PR   11-AUG-1999;   99US-0148319.
PR   12-AUG-1999;   99US-0148341.
PR   13-AUG-1999;   99US-0148565.
```

```
PR   13-AUG-1999;   99US-0148684.
PR   16-AUG-1999;   99US-0149368.
PR   17-AUG-1999;   99US-0149175.
PR   18-AUG-1999;   99US-0149426.
PR   20-AUG-1999;   99US-0149722.
PR   20-AUG-1999;   99US-0149723.
PR   20-AUG-1999;   99US-0149929.
PR   23-AUG-1999;   99US-0149902.
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match            99.6%;  Score 2535;  DB 21;  Length 471;
  Best Local Similarity  99.6%;  Pred. No. 3.5e-262;
  Matches  469;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 MEADESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRR 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 meadesgislpsgpdgrkrrvsyfyeptigdyyygqghpmkphrirmahsliihyhlhrr 60

Qy      61 LEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFD 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 leisrpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrfnvgedcpvfdglfd 120

Qy     121 FCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFK 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 fcrasaggsigaavklnrqdadiainwggglhhakkseasgfcyvndivlgilellkmfk 180

Qy     181 RVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPL 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 rvlyididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdvgaekgkyyalnvpl 240

Qy     241 NDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFL 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ndgmddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcfnlsvkghadclrfl 300

Qy     301 RSYNVPLMVLGGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPS 360
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 rsynvplmvlggggytirnvarcwcyetavavgvepdnklpyneyfeyfgpdytlhvdps 360

Qy     361 PMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWS 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db     361 pmenlntpkdmerirntlleqlsglihapsvqfqhtppvnrvldepeddmetrpkpriws 420

Qy     421 GTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDVNPPSS 471
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gtatyesdsddddkplhgyscrggattdrdstgedemdddnpepdvnppss 471


RESULT   4
AAG05188
ID   AAG05188 standard; Protein; 432 AA.
XX
AC   AAG05188;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 1500.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
```

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

```
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           91.1%;  Score 2318;  DB 21;  Length 432;
  Best Local Similarity  99.5%;  Pred. No. 5.1e-239;
  Matches  430;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      40 MKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSA 99
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mkphrirmahsliihyhlhrrleisrpsladasdigrfhspeyvdflasvspesmgdpsa 60

Qy     100 ARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEA 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 arnlrrfnvgedcpvfdglfdfcrasaggsigaavklnrqdadiainwgggLhhakksea 120

Qy     160 SGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPG 219
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 sgfcyvndivlgilellkmfkrvlyididvhhgdgveeafyttdrvmtvsfhkfgdffpg 180

Qy     220 TGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLS 279
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tghirdvgaekgkyyalnvplndgmddesfrslfrpliqkvmevyqpeavvlqcgadsls 240

Qy     280 GDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNK 339
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     241 gdrlgcfnlsvkghadclrflrsynvplmvlggggytirnvarcwcyetavavgvepdnk 300

Qy     340 LPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPV 399
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 lpyneyfeyfgpdytlhvdpspmenlntpkdmerirntlleqlsglihapsvqfqhtppv 360

Qy     400 NRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDD 459
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     361 nrvldepeddmetrpkpriwsgtatyesdsddddkplhgyscrggattdrdstgedemdd 420

Qy     460 DNPEPDVNPPSS 471
           ||||||||||||
Db     421 dnpepdvnppss 432


RESULT   5
AAG39373
ID   AAG39373 standard; Protein; 432 AA.
XX
AC   AAG39373;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 48706.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
PR   18-JUN-1999;   99US-0139457.
PR   18-JUN-1999;   99US-0139458.
PR   18-JUN-1999;   99US-0139459.
PR   18-JUN-1999;   99US-0139460.
PR   18-JUN-1999;   99US-0139461.
PR   18-JUN-1999;   99US-0139462.
PR   18-JUN-1999;   99US-0139463.
PR   18-JUN-1999;   99US-0139750.
PR   18-JUN-1999;   99US-0139763.
PR   21-JUN-1999;   99US-0139817.
PR   22-JUN-1999;   99US-0139899.
PR   23-JUN-1999;   99US-0140353.
PR   23-JUN-1999;   99US-0140354.
PR   24-JUN-1999;   99US-0140695.
PR   28-JUN-1999;   99US-0140823.
PR   29-JUN-1999;   99US-0140991.
PR   30-JUN-1999;   99US-0141287.
PR   01-JUL-1999;   99US-0141842.
PR   01-JUL-1999;   99US-0142154.
PR   02-JUL-1999;   99US-0142055.
PR   06-JUL-1999;   99US-0142390.
PR   08-JUL-1999;   99US-0142803.
PR   09-JUL-1999;   99US-0142920.
PR   12-JUL-1999;   99US-0142977.
PR   13-JUL-1999;   99US-0143542.
PR   14-JUL-1999;   99US-0143624.
PR   15-JUL-1999;   99US-0144005.
```

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 91.1%; Score 2318; DB 21; Length 432;
Best Local Similarity 99.5%; Pred. No. 5.1e-239;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      40 MKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSA 99
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mkphrirmahsliihyhlhrrleisrpsladasdigrfhspeyvdflasvspesmgdpsa 60

Qy     100 ARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEA 159
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 arnlrrfnvgedcpvfdglfdfcrasaggsigaavklnrqdadiainwgggLhhakksea 120

Qy     160 SGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPG 219
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 sgfcyvndivlgilellkmfkrvlyididvhhgdgveeafyttdrvmtvsfhkfgdffpg 180

Qy     220 TGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLS 279
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tghirdvgaekgkyyalnvplndgmddesfrslfrpliqkvmevyqpeavvlqcgadsls 240

Qy     280 GDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNK 339
           |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     241 gdrlgcfnlsvkghadclrflrsynvplmvlggggytirnvarcwcyetavavgvepdnk 300

Qy     340 LPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPV 399
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 lpyneyfeyfgpdytlhvdpspmenlntpkdmerirntlleqlsglihapsvqfqhtppv 360

Qy     400 NRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDD 459
           |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     361 nrvldepeddmetrpkpriwsgtatyesdsddddkplhgyscrggattdrdstgedemdd 420

Qy     460 DNPEPDVNPPSS 471
           ||||||||||||
Db     421 dnpepdvnppss 432
```

RESULT 6
AAG05189
ID AAG05189 standard; Protein; 425 AA.
XX

AC AAG05189;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 1501.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

```
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match           89.6%;  Score 2279;  DB 21;  Length 425;
  Best Local Similarity  99.5%;  Pred. No. 7.4e-235;
  Matches  423;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      47 MAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRF 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mahsliihyhlhrrleisrpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrf 60

Qy     107 NVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVN 166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 nvgedcpvfdglfdfcrasaggsigaavklnrqdadiainwggglhhakkseasgfcyvn 120

Qy     167 DIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDV 226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 divlgilellkmfkrvlyididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdv 180

Qy     227 GAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCF 286
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gaekgkyyalnvplndgmddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcf 240

Qy     287 NLSVKGHADCLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYF 346
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     241 nlsvkghadclrflrsynvplmvlggggytirnvarcwcyetavavgvepdnklpyneyf 300

Qy     347 EYFGPDYTLHVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEP 406
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 eyfgpdytlhvdpspmenlntpkdmerirntlleqlsglihapsvqfqhtppvnrvldep 360

Qy     407 EDDMETRPKPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDV 466
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db     361 eddmetrpkpriwsgtatyesdsddddkplhgyscrggattdrdstgedemdddnpepdv 420

Qy     467 NPPSS 471
           |||||
Db     421 nppss 425


RESULT   7
AAG39374
ID   AAG39374 standard; Protein; 425 AA.
XX
AC   AAG39374;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 48707.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
```

PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```
  Query Match             89.6%;  Score 2279;  DB 21;  Length 425;
  Best Local Similarity   99.5%;  Pred. No. 7.4e-235;
  Matches  423;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      47 MAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRF 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 mahsliihyhlhrrleisrpsladasdigrfhspeyvdflasvspesmgdpsaarnlrrf 60

Qy     107 NVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVN 166
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 nvgedcpvfdglfdfcrasaggsigaavklnrqdadiainwgggLhhakkseasgfcyvn 120

Qy     167 DIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDV 226
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 divlgilellkmfkrvlyididvhhgdgveeafyttdrvmtvsfhkfgdffpgtghirdv 180

Qy     227 GAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCF 286
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gaekgkyyalnvplndgmddesfrslfrpliqkvmevyqpeavvlqcgadslsgdrlgcf 240

Qy     287 NLSVKGHADCLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYF 346
           |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     241 nlsvkghadclrflrsynvplmvlggggytirnvarcwcyetavavgvepdnklpyneyf 300

Qy     347 EYFGPDYTLHVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEP 406
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 eyfgpdytlhvdpspmenlntpkdmerirntlleqlsglihapsvqfqhtppvnrvldep 360

Qy     407 EDDMETRPKPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDV 466
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 eddmetrpkpriwsgtatyesdsddddkplhgyscrggattdrdstgedemdddnpepdv 420

Qy     467 NPPSS 471
           |||||
Db     421 nppss 425


RESULT   8
AAY28797
ID   AAY28797 standard; protein; 458 AA.
XX
AC   AAY28797;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-1.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90837.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 49-51; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   458 AA;


  Query Match             68.2%;  Score 1734;  DB 20;  Length 458;
  Best Local Similarity   69.3%;  Pred. No. 1.8e-176;
  Matches  328;  Conservative   50;  Mismatches   73;  Indels   22;  Gaps    7;

Qy       1 MEADESGISLPS---GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHL 57
           | |   | ||||   | |  :|||||||||:|||||||||||||||||||||||::|| |
Db       1 maasgegaslpspaggedahrrrvsyfyepsigdyyyggghpmkphrirmahslvvhygl 60

Qy      58 HRRLEISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSA--ARNLRRFNVGEDCPVF 115
           || ||:|||  |  :|| |||| :|| |||| :    |:|     | ::||||||||||
Db      61 hrllelsrpypaseadirrfhsddyvaflasat----gnpgvldpraikrfnvgedcpvf 116

Qy     116 DGLFDFCRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILEL 175
           |||| ||:|||||||||||||| |||| :|| ||||||||||||||||||||| ||||
Db     117 dglfpfcqasaggsigaavklnrgdaditvnwagglhhakkseasgfcyvndivlailel 176

Qy     176 LKMFKRVLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYA 235
           ||  :||||:|||||||||||||:||:||||||||:|||||||| |||| :||:||
Db     177 lkfhrrvlyvdidvhhgdgveeaffttnrvmtvsfhkygdffpgtghitdvgaaegkhya 236

Qy     236 LNVPLNDGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHAD 295
           |||||:||:|| :|| ||: :|:|||||||| ||||||||||:|||||||||||||||||
Db     237 invplsdgiddttfrglfqciikkvmevyqpdvvvlqcgadslagdrlgcfnlsvkghad 296

Qy     296 CLRFLRSYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTL 355
           ||||||||||:||||| ||||||||||||||||||||||||||||:|:||||||||||
Db     297 clrflrsynvpmmvlggggytirnvarcwcyetavavgvepdnklpyndyyeyfgpdytl 356

Qy     356 HVDPSPMENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRP- 414
           |: |  :||||| ||:| |:| :|| || : | || ||   |      :| |:||: ||
Db     357 hiqpksvenlnttkdlenikamilenlskiehvpstqfhdrpsdpeapeekeedmdkrpp 416

Qy     415 -KPRXWSGTATYESDSDDDDKPLHGYSCRGGATTDRDSTGEDEMDDDNPEPDV 466
            : | ||| | |:||::| |            :  :| |   :| |: |: |:
Db     417 qrsrlwsgga-ydsdtedpdsl---------ksegkdvtanfqmkde-pkddl 458


RESULT   9
AAB67811
ID   AAB67811 standard; Protein; 501 AA.
XX
AC   AAB67811;
XX
DT   29-JUN-2001  (first entry)
XX
DE   Amino acid sequence of a histone deacetylase designated AtRPD3A.
XX
KW   Histone deacetylase; AtRPD3A; RPD3; gene expression; transgenic plant;
KW   HDA1; ethylene-responsive phenotype; hypocotyl elongation.
XX
OS   Arabidopsis thaliana.
XX
PN   CA2316036-A1.
XX
PD   27-FEB-2001.
XX
PF   24-AUG-2000; 2000CA-2316036.
XX
```

PR 27-AUG-1999; 99US-0383971.
XX
PA (MIAC) CANADA MIN AGRIC & AGRI-FOOD CANADA.
XX
PI Miki B, Brown D, Tian L, Wu K;
XX
DR WPI; 2001-258457/27.
DR N-PSDB; AAF80350.
XX
PT Methods for regulating gene expression in transgenic plants, e.g.
PT repressing ethylene-responsive phenotypes (e.g. inhibition of hypocotyl
PT elongation), comprises introducing genes encoding histone deacetylase
PT -
XX
PS Claim 10; Fig 1A; 91pp; English.
XX
CC The present sequence represents Arabidopsis thaliana histone deacetylase
CC designated AtRPD3A. The protein is homologous to yeast RPD3 and HDA1.
CC The polynucleotide sequence is used in the method of the invention.
CC The specification describes a method for regulating gene expression in
CC transgenic plants. The method comprises modifying histones by introducing
CC chimeric nucleotide sequences which have regulatory elements in operative
CC association with a gene of interest or with a nucleotide sequence
CC encoding histone deacetylase. The method is useful for regulating the
CC developmental, physiological or biochemical pathway within a plant,
CC particularly for repressing ethylene-responsive phenotypes
CC (e.g. inhibition of hypocotyl elongation). The method is also useful
CC as a functional test for identifying a phenotype associated with
CC perturbing a gene. The histone deacetylase genes are useful for
CC altering the development of an organism.
XX
SQ Sequence 501 AA;

Query Match 61.4%; Score 1561.5; DB 22; Length 501;
Best Local Similarity 60.7%; Pred. No. 5.8e-158;
Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;

```
Qy       4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
           |  | || ||||| ||:| |||:| :|:|||||||||||||||| |:|: || | : :::
Db       2 dtggnslasgpdgvkrkvcyfydpevgnyyyggqhpmkphrirmthallahygllqhmqv 61

Qy      64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
             :|  |   |: |||: :|| || |::||:  |    | |:||||||||||||: ||:
Db      62 lkpfparerdlcrfhaddyvsflrsitpetqqd--qirqlkrfnvgedcpvfdglysfcq 119

Qy     124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
             ||||:| :|||    |||||| |||||||| |||||||||||| ||||||  :|||
Db     120 tyaggsvggsvklnhgicdiainwagglhhakkceasgfcyvndivlailellkqhervl 179

Qy     184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
           |:|||:|||||||||| ||||||||||||||:|||||||:|:|   ||||:|||||:||
Db     180 yvdidihhgdgveeafyatdrvmtvsfhkfgdyfpgtghiqdigygsgkyyslnvplddg 239

Qy     244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
           :|||||:  ||:|:: ||||:::| ||||||||||||||||||||||:||||:|::|:||:
Db     240 iddesyhllfkpimgkvmeifrpgavvlqcgadslsgdrlgcfnlsikghaecvkfmrsf 299

Qy     304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363
           ||||::||| |||||||||||||| ||:||| ::|:| :||:|||||||||| || ||
Db     300 nvpllllggggytirnvarcwcyetgvalgvevedkmpeheyyeyfgpdytlhvapsnme 359

Qy     364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXWSG 421
           | |: : :| ||| ||  || | |||||| ||  ||      :|| ::| : |  |
Db     360 nknsrqmleeirndllhnlsklqhapsvpfqerppdtetpevdedqedgdkrwdpds--- 416

Qy     422 TATYESDSDDDDKP------------------LHGYSCRG-GATTDRDSTGEDEMDDDN 461
               : | ||| ||                  | |   || |   : | :|  ::   |
Db     417 ----dmdvdddrkpipsrvkreavepdtkdkdglkgimergkgcevevdesgstkvtgvn 472

Qy     462 P 462
           |
Db     473 p 473
```

RESULT 10
AAB67165
ID AAB67165 standard; Protein; 501 AA.
XX
AC AAB67165;
XX
DT 12-APR-2001 (first entry)
XX
DE Arabidopsis histone deacetylase 1 HDAC1.
XX
KW Gene expression; disease model; treatment; suppression;
KW chromatin inactivation; DNA binding site; histone deacetylation.
XX
OS Arabidopsis thaliana.
XX
PN WO200102019-A2.
XX
PD 11-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-GB02497.
XX
PR 30-JUN-1999; 99GB-0015126.
XX
PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX
PI Buluwela L, Ali S;
XX
DR WPI; 2001-138069/14.
XX
PT Suppressing expression of selected gene for treating cancer, involves
PT introducing peptide comprising nucleic acid binding portion that binds
PT to or associated with selected gene and chromatin inactivation portion
PT -
XX
PS Disclosure; Fig 5; 65pp; English.
XX
CC The present invention describes a method of suppressing the expression of
CC a gene of interest by introducing into the cell a protein containing a
CC DNA binding site and a chromatin inactivation portion, or a nucleotide
CC encoding such a peptide. Preferably the chromatin inactivation portion is
CC part of the histone deacetylation (HDAC) complex. The method can be used
CC in disease treatment, for example in the treatment of cancer by the
CC suppression of oncogenes, and in the production of disease models.
XX
SQ Sequence 501 AA;

Query Match 61.4%; Score 1561.5; DB 22; Length 501;
Best Local Similarity 60.7%; Pred. No. 5.8e-158;
Matches 292; Conservative 64; Mismatches 94; Indels 31; Gaps 5;

```
Qy       4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
           |  | || ||||| ||:| |||:| :|:|||||||||||||||| |:|: || | : :::
Db       2 dtggnslasgpdgvkrkvcyfydpevgnyyyggqhpmkphrirmthallahygllqhmqv 61

Qy      64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
             :|  |   |: |||: :|| || |::||:  |    | |:||||||||||||: ||:
Db      62 lkpfparerdlcrfhaddyvsflrsitpetqqd--qirqlkrfnvgedcpvfdglysfcq 119

Qy     124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
             ||||:| :|||    |||||| |||||||| |||||||||||| ||||||  :|||
Db     120 tyaggsvggsvklnhglcdiainwagglhhakkceasgfcyvndivlailellkqhervl 179

Qy     184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
           |:|||:|||||||||| ||||||||||||||:|||||||:|:|   ||||:|||||:||
Db     180 yvdidihhgdgveeafyatdrvmtvsfhkfgdyfpgtghiqdigygsgkyyslnvplddg 239

Qy     244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
           :|||||:  ||:|:: ||||:::| ||||||||||||||||||||||:||||:|::|:||:
Db     240 iddesyhllfkpimgkvmeifrpgavvlqcgadslsgdrlgcfnlsikghaecvkfmrsf 299
```

```
Qy    304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363
          ||||::||| |||||||||||||||| ||:||| ::|:| :||:|||||||||| || ||
Db    300 nvpllllggggytirnvarcwcyetgvalgvevedkmpeheyyeyfgpdytlhvapsnme 359

Qy    364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXWSG 421
          | |: : :| ||| ||  || | ||||| ||  ||      :|| ::| : |  |
Db    360 nknsrqmleeirndllhnlsklqhapsvpfqerppdtetpevdedqedgdkrwdpds--- 416

Qy    422 TATYESDSDDDDKP-----------------LHGYSCRG-GATTDRDSTGEDEMDDDN 461
              : | ||| ||                  | |   || |   : | :|  ::   |
Db    417 ----dmdvdddrkpipsrvkreavepdtkdkdglkgimergkgcevevdesgstkvtgvn 472

Qy    462 P 462
          |
Db    473 p 473
```

RESULT 11
AAY58829
ID AAY58829 standard; Protein; 473 AA.
XX
AC AAY58829;
XX
DT 08-MAY-2000 (first entry)
XX
DE Soybean histone deacetylase 1 (HD1) protein.
XX
KW Chromatin associated protein; histone deacetylase gene 1; HD1;
KW soybean; transgenic plant; transcription regulation.
XX
OS Glycine max.
XX
PN WO200004177-A1.
XX
PD 27-JAN-2000.
XX
PF 13-JUL-1999; 99WO-US15807.
XX
PR 14-JUL-1998; 98US-0092841.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Cahoon RE, Vollmer SJ;
XX
DR WPI; 2000-182439/16.
DR N-PSDB; AAZ58260.
XX
PT New nucleic acid fragment useful as probes and primers, for
PT transforming plants -
XX
PS Claim 1; Page 27-28; 36pp; English.
XX
CC The present sequence is that of soybean histone deacetylase 1
CC (HD1), a chromatin associated protein, as deduced from a soybean
CC root cDNA clone (see AAZ58260) isolated on the basis of homology to
CC plant histone deacetylases. The invention relates to isolated
CC rice, soybean and wheat nucleic acid fragments encoding HD1. It
CC also relates to the construction of a chimeric gene encoding all or
CC a portion of HD1, in sense or antisense orientation, where
CC expression of the chimeric gene results in production of altered
CC levels of HD1 in a transformed host cell. The availability of
CC nucleic acid sequences encoding (portions) of histone deacetylase
CC proteins will facilitate studies of global transcriptional
CC regulation in eukaryotic cells, and will also provide mechanisms to
CC control transcriptional gene regulation in plants.
XX
SQ Sequence 473 AA;

Query Match 61.3%; Score 1558.5; DB 21; Length 473;
Best Local Similarity 61.8%; Pred. No. 1.1e-157;
Matches 290; Conservative 71; Mismatches 83; Indels 25; Gaps 7;

```
Qy      4 DESGISLPSGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEI 63
          :  | ||||| || ||:|||||:| :|:||||||||||||||| |:|: || | : :::
Db      2 esggnslpsgsdgvkrkvsyfydpevgnyyygqghpmkphrirmthallahygllqhmqv 61

Qy     64 SRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCR 123
           :|  |   |: :||: :|| ||  ::||:  |    [ |:||||||||||||||: ||:
Db     62 lkpmaakdrdlckfhaddyvaflrgitpetqqd--qlrqlkrfnvgedcpvfdglysfcq 119

Qy    124 ASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVL 183
            ||||:| |:|||    |||||| |||||||| |||||||||||| |||||: :|||
Db    120 tyaggsvggalklnhgvcdiainwagglhhakkceasgfcyvndivlailellkihervl 179

Qy    184 YIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDG 243
          |:|||:|||||||||||||||||||||||:||||| |||:|  |||||:|||||:||
Db    180 yvdidihhgdgveeafyttdrvmtvsfhkfgdyfpgtgdirdigyakgkyyslnvpldda 239

Qy    244 MDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSY 303
          :||||::|||:|:: ||||:::| ||||||||||||||||||||:||||:|:|::||:
Db    240 iddesyqslfkpimgkvmeifrpgavvlqcgadslsgdrlgcfnlsikghaecvrymrsf 299

Qy    304 NVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPME 363
          ||||::||| |||||||||||:||:||:|:| |:|:| :||:|||||||||| || ||
Db    300 nvpllliggggytirnvarcwcfetsvalgielddkmpqheyyeyfgpdytlhvapsnme 359

Qy    364 NLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVL--DEPEDDMETRPKPRXWSG 421
          | |: : :: ||  ||: || | ||||| ||  ||   :|  || :|| : |
Db    360 nknsrqlldeiraklldnlsrlqhapsvpfqerppdaellerdedqddrder-------- 411

Qy    422 TATYESDSD----DDDKPLH---GYSCRGGATTDRDS-TGEDEMDDDNP 462
             :: |||    ||  |:       |      |:|: :| | |  | |
Db    412 ---wdpdsdrevgddsnpvrrrvksecvd--aedkdtvsgvdsmavdep 455
```

RESULT 12
AAY28800
ID AAY28800 standard; protein; 517 AA.
XX
AC AAY28800;
XX
DT 13-JAN-2000 (first entry)
XX
DE Maize histone deacetylase-4.
XX
KW Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;
KW acetyl modification; promoter; regulatory element; transgenic plant;
KW disease resistance; toxin screening; pathogenicity;
KW disease response promoter.
XX
OS Zea mays.
XX
PN WO9951731-A2.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07370.
XX
PR 03-APR-1998; 98US-0080563.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Baldwin DA, Briggs SP, Crane VC;
XX
DR WPI; 1999-611038/52.
DR N-PSDB; AAX90840.
XX
PT New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance -
XX
PS Claim 1; Page 63-65; 87pp; English.
XX
CC The present sequence is maize histone deacetylase encoded by HD cDNA
CC belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl
CC modifications, may be localised to promoters targeted for repression by

```
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   517 AA;


  Query Match             61.0%;  Score 1551.5;  DB 20;  Length 517;
  Best Local Similarity   66.4%;  Pred. No. 7.2e-157;
  Matches  288;  Conservative   59;  Mismatches   72;  Indels   15;  Gaps    5;

Qy       3 ADESGISLPS-GPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
           |   | ||||  ||||:|||| |||:| :|:|||||||||||||||||| |||:  | |  ::
Db       6 agsggnslpsvgpdgqkrrvcyfydpdvgnyyyyqqghpmkphrirmthsllarygllnqm 65

Qy      62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDF 121
           :: ||: |   |: |||: :|::|| ||:||:  |    | |:|||||||||||||: |
Db      66 qvyrpnpardrdlcrfhaddyinflrsvtpetqqd--qirllkrfnvgedcpvfdglysf 123

Qy     122 CRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKR 181
           |:  || |:| |||||    |||||| |||||||| |||||||||||| |||||  :|
Db     124 cqtyagasvggavkln-hghdiainwsgglhhakkceasgfcyvndivlailellkhher 182

Qy     182 VLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLN 241
           |||:|||:|||||||||||||||||||||||||:||||| ||||:|  |||||:|||||:
Db     183 vlyvdidihhgdgveeafyttdrvmtvsfhkfgdyfpgtgdirdighskgkyyslnvpld 242

Qy     242 DGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLR 301
           ||:||||::|||:|:: |||||::| |||||||||||||||||||||:||||:|:|:::|
Db     243 dgiddesyqslfkpimgkvmevfrpgavvlqcgadslsgdrlgcfnlsikghaecvrymr 302

Qy     302 SYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP 361
           |:||||::||| |||||||||||||| ||:| ||::|:| |||:||||||||| ||
Db     303 sfnvpllllggggytirnvarcwcyetgvalgqepedkmpvneyyeyfgpdytlhvapsn 362

Qy     362 MENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRV--LDEPEDDMETRPKPRXW 419
           ||| || : :: ||: ||: || | ||||| ||   |   :   || :|| : |  |
Db     363 menkntrqqlddirskllднlsklrhapsvhfqervpdteipeqdedqddpderhdp--- 419

Qy     420 SGTATYESDSDDDD 433
                 :|| : ||
Db     420 ------dsdmevdd 427


RESULT  13
AAY28799
ID   AAY28799 standard; protein; 439 AA.
XX
AC   AAY28799;
XX
DT   13-JAN-2000  (first entry)
XX
DE   Maize histone deacetylase-3.
XX
KW   Maize histone deacetylase; HD; HD cDNA; family 1, ZmHD1; gene repression;
KW   acetyl modification; promoter; regulatory element; transgenic plant;
KW   disease resistance; toxin screening; pathogenicity;
KW   disease response promoter.
XX
OS   Zea mays.
XX
PN   WO9951731-A2.
XX
PD   14-OCT-1999.
XX
PF   02-APR-1999;   99WO-US07370.
XX
PR   03-APR-1998;   98US-0080563.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Baldwin DA,  Briggs SP,  Crane VC;
XX
DR   WPI; 1999-611038/52.
DR   N-PSDB; AAX90839.
XX
PT   New deacetylase genes, used for producing transgenic plants which have
PT   increased disease resistance  -
XX
PS   Claim 1; Page 58-60; 87pp; English.
XX
CC   The present sequence is maize histone deacetylase encoded by HD cDNA
CC   belonging to family 1, ZmHD1. This enzyme responsible for removing acetyl
CC   modifications, may be localised to promoters targeted for repression by
CC   other proteins that associate with HD and specifically bind regulatory
CC   elements in promoter DNA. The HD nucleotide sequence can be used for
CC   producing transgenic plants with increased disease resistance.
CC   Additionally, compositions find use in screening for toxins that affect
CC   pathogenicity and in determining which disease response promoters are
CC   regulated by histone deacetylase.
XX
SQ   Sequence   439 AA;


  Query Match             59.7%;  Score 1520;  DB 20;  Length 439;
  Best Local Similarity   66.0%;  Pred. No. 1.3e-153;
  Matches  289;  Conservative   56;  Mismatches   77;  Indels   16;  Gaps    6;

Qy       3 ADESGISLP-SGPDGRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRL 61
           |   | ||| :| || |||| |||:  :|:|||||||||||||||||| |:|:  | |  ::
Db       5 aggggnslpttgadgskrrvcyfydaevgnyyyyqqghpmkphrirmthallgrygllдqm 64

Qy      62 EISRPSLADASDIGRFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDF 121
           :: ||  |   |: |||: :|| || ||:||:  |    | |:|||||||||||||: |
Db      65 qvfrphpardrdlcrfhaddyvsflrsvtpetqqd--qiralkrfnvgedcpvfdglysf 122

Qy     122 CRASAGGSIGAAVKLNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKR 181
           |:  |||||:| ||||||    |||||| |||||||| |||||||||||| |||||  :|
Db     123 cqtyaggsvggavkln-hghdiainwagglhhakkceasgfcyvndivlailellkyhqr 181

Qy     182 VLYIDIDVHHGDGVEEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLN 241
           ||||:|||:|||||||||||||||||||||||||:||||| |||||  |||||:|||||:
Db     182 vlyvdidihhgdgveeafyttdrvmtvsfhkfgdyfpgtgdirdvghskgkyyslnvpld 241

Qy     242 DGMDDESFRSLFRPLIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLR 301
           ||:|||||::|||:|:: ||||||: | |||||||||||||||||||:||||:|:|:|:|
Db     242 dgiddesyqslfkpimgkvmevfnpgavvlqcgadslsgdrlgcfnlsikghaecvrfmr 301

Qy     302 SYNVPLMVLGGEGYTIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSP 361
           |:||||::||| |||||||||||||| ||:| |  :|:| |||:||||||||| ||
Db     302 sfnvpllllggggytirnvarcwcyetgvalgheltdkmppneyyeyfgpdytlhvapsn 361

Qy     362 MENLNTPKDMERIRNTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSG 421
           ||| ||   :: |:: ||: || | ||||||||  ||   : :: ||  :  |  |
Db     362 menkntrhqlddiksklldnlsklrhapsvqfqerppeaelpeqded--kenpder---- 415

Qy     422 TATYESDSD---DDDKPL 436
              :::|||   :| |||
Db     416 ---hdadsdvemndakpl 430


RESULT  14
AAY58828
ID   AAY58828 standard; Protein; 493 AA.
XX
AC   AAY58828;
XX
DT   08-MAY-2000  (first entry)
XX
DE   Rice histone deacetylase 1 (HD1) protein (C-terminal portion).
XX
KW   Chromatin associated protein; histone deacetylase gene 1; HD1;
KW   rice; transgenic plant; transcription regulation.
```

```
XX
OS   Oryza sativa.
XX
PN   WO200004177-A1.
XX
PD   27-JAN-2000.
XX
PF   13-JUL-1999;   99WO-US15807.
XX
PR   14-JUL-1998;   98US-0092841.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Cahoon RE,  Vollmer SJ;
XX
DR   WPI; 2000-182439/16.
DR   N-PSDB; AAZ58259.
XX
PT   New nucleic acid fragment useful as probes and primers, for
PT   transforming plants  -
XX
PS   Claim 1; Page 25-26; 36pp; English.
XX
CC   The present sequence is that of a C-terminal portion of rice
CC   histone deacetylase 1 (HD1), a chromatin associated protein, as
CC   deduced from a cDNA contig (see AAZ58259) obtained from cDNA clones
CC   isolated from rice callus, 15-day-old leaf and 15-day-old seedling
CC   cDNA libraries on the basis of homology to other plant histone
CC   deacetylases.  The invention relates to isolated rice, soybean and
CC   wheat nucleic acid fragments encoding HD1.  It also relates to the
CC   construction of a chimeric gene encoding all or a portion of HD1,
CC   in sense or antisense orientation, where expression of the chimeric
CC   gene results in production of altered levels of HD1 in a
CC   transformed host cell.  The availability of nucleic acid sequences
CC   encoding (portions) of histone deacetylase proteins will facilitate
CC   studies of global transcriptional regulation in eukaryotic cells,
CC   and will also provide mechanisms to control transcriptional gene
CC   regulation in plants.
XX
SQ   Sequence   493 AA;


  Query Match             58.9%;  Score 1498;  DB 21;  Length 493;
  Best Local Similarity   64.4%;  Pred. No. 3.5e-151;
  Matches  286;  Conservative   55;  Mismatches   85;  Indels   18;  Gaps    7;

Qy      23 YFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIGRFHSPEY 82
           |||:  :|:|||||||||||||||| |:|: || |  :::: :|  |   |: |||: :|
Db       1 yfydaevgnyyyyqqghpmkphrirmthallahyglldqmqvlkphpardrdlcrfhaddy 60

Qy      83 VDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVKLNRQDAD 142
           | || ||:||:  |    | |:||||||||||||||: ||:  ||||:| |||||    |
Db      61 vaflrsvtpetqqd--qiralkrfnvgedcpvfdglysfcqtyaggsvggavkln-hghd 117

Qy     143 IAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGVEEAFYTT 202
           ||||| |||||||| |||||||||||| |||||  :||||:|||:|||||||||||||
Db     118 iainwagglhhakkceasgfcyvndivlailellkyhqrvlyvdidihhgdgveeafytt 177

Qy     203 DRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRPLIQKVME 262
           ||||||||||||:||||| |||:|  |||||:|||||:||:||||::|||:|:: ||||
Db     178 drvmtvsfhkfgdyfpgtgdirdighskgkyyslnvpldddgiddesyqslfkpimgkvme 237

Qy     263 VYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGYTIRNVAR 322
           |::| ||||||||||||||||||||::|||:|:||:||:||||::||| ||||||||
Db     238 vfrpgavvlqcgadslsgdrlgcfnlsirghaecvrfmrsfnvpllllqgggytirnvar 297

Qy     323 CWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIRNTLLEQL 382
           |||||| ||:| |  :|:| ||||||||||||||| || ||| || : :: ||: ||: |
Db     298 cwcyetgvalgheltdkmppneyfeyfgpdytlhvapsnmenkntrqqlddirsrlldnl 357

Qy     383 SGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSD---DDDKPLHGY 439
           | | |||||||  ||    | | ::| |  |  |       : :|||   || |||
Db     358 sklrhapsvqfqerpp-eaelpeqdedqed-pder-------hhadsdvemddvkpldds 408

Qy     440 SCRG---GATTDRDSTGEDEMDDD 460
               |         |:|   |  | |
Db     409 grrssiqnvrvkresaetdaadqd 432


RESULT  15
AAW29324
ID   AAW29324 standard; Protein; 482 AA.
XX
AC   AAW29324;
XX
DT   16-APR-1998  (first entry)
XX
DE   A novel histone deacetylase (HDx) protein, designated HD1.
XX
KW   Histone deacetylase gene; HDx; HD1; HDx polypeptide; deacetylation; H3;
KW   H4; cell differentiation; chromatin structure; cell cycle progression;
KW   proliferative disorder; fibroproliferative disorder;
KW   degenerative disorder; autoimmune disease; HDx inhibitor.
XX
OS   Homo sapiens.
XX
PN   WO9735990-A2.
XX
PD   02-OCT-1997.
XX
PF   26-MAR-1997;   97WO-US05275.
XX
PR   26-MAR-1996;   96US-0624735.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Hassig CA,  Jamison TF,  Schreiber SL,  Taunton J;
XX
DR   WPI; 1997-489651/45.
DR   N-PSDB; AAT86371.
XX
PT   New isolated histone deacetylase polypeptide(s) and genes - used to
PT   develop products for modulating the proliferation, survival or
PT   differentiation of cells, e.g. for treating tumours.
XX
PS   Claim 6; Pages 112-114; 159pp; English.
XX
CC   The present sequence represents a novel histone deacetylase (HDx)
CC   polypeptide, designated HD1. The HDx polypeptides are capable of
CC   modulating proliferation survival and differentiation of cells. The
CC   proteins are able to alter chromatin structure by deacetylating histones
CC   such as H3 or H4. They have the ability to modulate cell growth by
CC   influencing cell cycle progression or to modulate gene transcription. The
CC   products can be used for diagnosis and therapy. They can be used, for
CC   example, to treat tumours or proliferative disorders or spermatogenesis,
CC   osteogenesis, chondrogenesis or the differentiation of progenitor cells.
CC   They can also be used to treat psoriasis, bone diseases,
CC   fibroproliferative disorders, degenerative disorders, or for repair of
CC   cartilage, increasing bone density, liver repair subsequent to a partial
CC   hepatectomy, to promote regeneration of lung tissue in the treatment of
CC   emphysema, or for inducing tolerance in autoimmune diseases, and
CC   transplant recipients. HDx inhibitors can be used as anti-fungal agents,
CC   preservatives in foodstuff, feed supplements for promoting weight gain in
CC   livestock, disinfectants, insecticides or defoliants.  The products can
CC   also be used in cell cultures.
CC   note: in the claims, the full length HD1 sequence is incorrectly
CC   referred to as SEQ. ID. No.2; from the disclosure it is clear that HD1
CC   is SEQ. ID. No.5.
XX
SQ   Sequence   482 AA;


  Query Match             57.5%;  Score 1462.5;  DB 18;  Length 482;
  Best Local Similarity   58.5%;  Pred. No. 2.2e-147;
  Matches  262;  Conservative   79;  Mismatches   80;  Indels   27;  Gaps    4;
```

```
Qy      16 GRKRRVSYFYEPTIGDYYYGQGHPMKPHRIRMAHSLIIHYHLHRRLEISRPSLADASDIG 75
Db       6 gtrrkvcyyydgdvgnyyygqghpmkphrirmthnlllnyglyrkmeiyrphkanaeemt 65

Qy      76 RFHSPEYVDFLASVSPESMGDPSAARNLRRFNVGEDCPVFDGLFDFCRASAGGSIGAAVK 135
Db      66 kyhsddyikflrsirpdnmseys--kqmqrfnvgedcpvfdglfefcqlstggsvasavk 123

Qy     136 LNRQDADIAINWGGGLHHAKKSEASGFCYVNDIVLGILELLKMFKRVLYIDIDVHHGDGV 195
Db     124 lnkqqtdiavnwagglhhakkseasgfcyvndivlailellkyhqrvlyididihhgdgv 183

Qy     196 EEAFYTTDRVMTVSFHKFGDFFPGTGHIRDVGAEKGKYYALNVPLNDGMDDESFRSLFRP 255
Db     184 eeafyttdrvmtvsfhkygeyfpgtgdlrdigagkgkyyavnyplrdgiddesyeaifkp 243

Qy     256 LIQKVMEVYQPEAVVLQCGADSLSGDRLGCFNLSVKGHADCLRFLRSYNVPLMVLGGEGY 315
Db     244 vmskvmemfqpsavvlqcgsdslsgdrlgcfnltikghakcvefvksfnlpmlmlggggy 303

Qy     316 TIRNVARCWCYETAVAVGVEPDNKLPYNEYFEYFGPDYTLHVDPSPMENLNTPKDMERIR 375
Db     304 tirnvarcwtyetavaldteipnelpyndyfeyfgpdfklhispsnmtnqntneylekik 363

Qy     376 NTLLEQLSGLIHAPSVQFQHTPPVNRVLDEPEDDMETRPKPRXWSGTATYESDSDDDDKP 435
Db     364 qrlfenlrmlphapgvqmqaipe-daipeesgde----------------deddpdkr 404

Qy     436 LHGYSCRGGATTDRDSTGEDEMDDDNPE 463
Db     405 i------sicssdkriaceeefsdseee 426
```

Search completed: April 26, 2002, 17:52:14
Job time: 82 sec